



Search Engines Revealed

1. Introduction to Search Engine Optimization	4
1.1 SEO plan and objective	4
1.2 Defining the SEO budget	6
1.3 Avenues for promotion	6
2. Search Engines and Directories	8
2.1 How do Search Engines Work?	8
2.2 Human powered directories	11
2.3 Hybrid Search Engines	12
2.4 How do Search Engines rank web pages?	12
2.5 Most popular Search Engines	14
3. How major Search Engines work – crawling policies and algorithms	21
3.1 How Google Works	21
3.2 How AltaVista Works	37
3.3 How Teoma Works	49
3.4 How Inktomi Works	56
4. Introduction to Search Engine Submission	67
4.1 Search Engine Submission – Getting the Basics Right	67
4.2 The Submission Budget	69
4.3 Submitting to Google	70
4.4 Submitting to Inktomi	71
4.5 Submitting to Fast Search	71
4.6 Submitting to Teoma	72
4.7 Submitting to AltaVista	72

How You Can Finally Achieve Top Search Engine Rankings!

4.8 Submitting to Directories	75
4.9 Submitting to the big ones	79
5. Intelligent ways of improving page rankings.....	81
5.1 Link Popularity	82
5.2 Meta tags and keywords.....	88
5.3 Write a persuasive home page	95
5.4 Using Clickthroughs and Affiliate Programs for better ranking	102
5.5 Doorblocks	102
5.6 Submit a “site map” page or a page with links to all inner pages.....	103
5.7 Submit only key pages	104
5.8 Working with Frames.....	104
5.9 Using Robots.txt to your advantage.....	106
5.10 The Right Text Usage.....	109
5.11 Search Engine Regionalization.....	114
5.12 Spamming – A Definite No-No.....	116
5.13 Search Engine Cloaking explained.....	117
5.14 Does separating Domain Name Keywords with hyphens affect rankings?	119
5.15 Taking advantage of Alphanumeric Search Engine listings.....	120
5.16 Click Popularity.....	121
5.17 Registering a keyword rich domain.....	122
5.18 Web Hosting	124
5.19 Caps or no caps?	128
5.20 A few Search Engine myths	129
5.21 Things to Avoid.....	131
6. Paid Placement Programs	136
6.1 Google Adwords	137
6.2 Google AdSense	139
6.3 Overture	143
6.4 Searchfeed	146
6.5 Ah-Ha	147
6.6 Sprinks	148
6.7 Gaining more value out of pay-per-click programs	149
6.8 Disadvantages of Paid Programs	153
7. Verify and monitor listing and ranking.....	156
7.1 Spider spotting.....	156
7.2 URL Search.....	157
8. Measuring Performance – Web logs analysis	159

9. Review of SEO software and submission services.....162
10. ROI in search engine optimization163
11. Resources and Tools for Search Engine Marketing169
 11.1 Search Engine Marketing Professional Organization (SEMPO)169
 11.2 Tools.....170
 11.3 Search Engine Information Resources174

1. Introduction to Search Engine Optimization

A search engine is the most effective tool that can bring a prospective customer to your Company website. Millions of web visits are initiated daily through one or the other search engine to locate information or sources of supply. This is considered to be the most effective and targeted channel for you as a website owner to acquire a hot lead. Businesses all the over the world spend a huge sum on designing, building, maintaining and promoting their websites.

1.1 SEO plan and objective

Online advertising and marketing budgets have also soared. Relative to these, the investment required for getting traffic through search engines is much lower. However, as search engines have millions of pages in their coverage, it is important to have a proper approach to using this channel effectively.

The art and science of understanding how search engines identify pages that are relevant to a query made by a visitor and designing marketing strategies based on this is called search engine optimization. Search engines offer the most cost effective mechanism to acquire “real” and “live” business leads. It is found that in most cases search engine optimization delivers better ROI than other forms such as online advertisements, e-mail

marketing and newsletters, affiliate and pay per click advertising, and digital campaigns and promotions.

Before you begin to develop your own strategy and implementation plan to optimize your website for improved page ranking, you have to reflect on what is the major objective of this initiative. Is it your objective to attract more visitors to your site or convert more persons from being a visitor to a loyal stakeholder in your business? The dot com mania period showed a marked change in evaluating a website's success in terms of number of eyeballs that it could collect, without taking into perspective what these eyeballs or site visits meant to the organization. That has changed by now, and most companies realize that critical to a business unit's success is sticky visitors, loyalty and ultimately the impact this has on its bottom line.

Valuing a website in terms of what advertising it can attract and sustain is no longer the benchmark or performance indicator; what is certainly more important is what the site does to acquire and retain new and existing customers.

As search engines have several million pages that are available on the world-wide web, it is necessary to use specialized techniques to match your web page with the algorithms and ranking criteria that such engines use, thereby improving the chance of catching the limited attention span of the visitor.

The first step that advertisers and marketing professionals need to take to apply optimization techniques to a website is to articulate the objective and characterize the visitor, the desired visitor experience and outcome. The optimization plan should evolve out of this.

1.2 Defining the SEO budget

The word optimization is by itself suggestive that the plan should balance the initiative and the budget so as to get cost effective results. If the stakes are high, it may sustain higher advertisement and paid listing options. How much would you like to spend on this exercise? If the budget is limited, the expensive options of several advertisements, linking programs, directory listings will have to be forsaken and attention given to getting the best results from limited but focused efforts. The key metrics to this program is to assess the ROI it delivers: marketing dollars vis-à-vis measurable benefits to the organization.

1.3 Avenues for promotion

Search Engines provide you an effective vehicle for promotion of your website. There are no doubt other channels available. 'Directories' is often included in the generic term 'search engines' although they are distinct in their characteristics and function. Advertisements through banner displays at popular and often visited sites and "portals", reciprocal links, affiliate links and programs that direct visitors from one site to a targeted

site (with a payment associated with such arrangements) and publicity through other media are other well known avenues of promotion. Mass email campaigns, publishing and distributing internet newsletters, 'permission marketing' using list servers and internet based marketing promotions including coupons and sweepstakes are other forms of online advertising. The key determinant is what value the initiative offers you in return for the investment. Metrics are somewhat more difficult to establish for search engine optimization compared to other direct forms of advertisement. Some suggested measures for determining ROI are discussed in a later section.

2. Search Engines and Directories

You would be using search engines so you know how they work from the user perspective. From your own experience as a user, you also know that only those results that list at the top of the heap are most likely to attract you. It doesn't amuse you to know that your search yielded 44316 results. Perhaps even number 50 on your list will not get your custom or even your attention. Thus you know that getting listed on the top or as near to the top is crucial. Since most of the search engine traffic is free, you'll usually find it worth your time to learn a few tricks to maximize the results from your time and effort. In the next section, you will see how search engine works – from your perspective as a website owner.

2.1 How do Search Engines Work?

Most of the top-ranked search engines are crawler based search engines while some may be based on human compiled directories. The people behind the search engines want the same thing every webmaster wants - traffic to their site. Since their content is mainly links to other sites, the thing for them to do is to make their search engine bring up the most relevant sites to the search query, and to display the best of these results first.

In order to accomplish this, they use a complex set of rules called algorithms. When a search query is submitted at a search engine, sites are determined to be relevant or not

relevant to the search query according to these algorithms, and then ranked in the order it calculates from these algorithms to be the best matches first.

Search engines keep their algorithms secret and change them often in order to prevent webmasters from manipulating their databases and dominating search results. They also want to provide new sites at the top of the search results on a regular basis rather than always having the same old sites show up month after month.

An important difference to realize is that search engines and directories are not the same. Search engines use a spider to "crawl" the web and the web sites they find, as well as submitted sites. As they crawl the web, they gather the information that is used by their algorithms in order to rank your site.

Directories rely on submissions from webmasters, with live humans viewing your site to determine if it will be accepted. If accepted, directories often rank sites in alphanumeric order, with paid listings sometimes on top. Some search engines also place paid listings at the top, so it's not always possible to get a ranking in the top three or more places unless you're willing to pay for it.

Let us now look at a more detailed explanation on how Search Engines work. Crawler based search engines are primarily composed of three parts.

2.1.1 Spidering

A search engine robot's action is called spidering, as it resembles the multiple legged spiders. The spider's job is to go to a web page, read the contents, connect to any other pages on that web site through links, and bring back the information. From one page it will travel to several pages and this proliferation follows several parallel and nested paths simultaneously. Spiders frequent the site at some interval, may be a month to a few months, and re-index the pages. This way any changes that may have occurred in your pages could also be reflected in the index. The spiders automatically visit your web pages and create their listings. An important aspect is to study what factors promote "deep crawl" – the depth to which the spider will go into your website from the page it first visited. Listing (submitting or registering) with a search engine is a step that could accelerate and increase the chances of that engine "spidering" your pages.

2.1.2 Indexing

The spider's movement across web pages stores those pages in its memory, but the key action is in indexing. The index is a huge database containing all the information brought back by the spider. The index is constantly being updated as the spider collects more information. The entire page is not indexed and the searching and page-ranking algorithm is applied only to the index that has been created. Most search engines claim that they index the full visible body text of a page. In a subsequent section, we explain the key considerations to ensure that indexing of your web pages improves relevance during search. The combined understanding of the indexing and the page-ranking process will lead to developing the right strategies. The Meta tags 'Description' and 'Keywords' have

a vital role as they are indexed in a specific way. Some of the top search engines do not index the keywords that they consider spam. They will also not index certain ‘stop words’ (commonly used words such as ‘a’ or ‘the’ or ‘of’) so as to save space or speed up the process. Images are obviously not indexed, but image descriptions or Alt text or “text within comments” is included in the index by some search engines.

2.1.3 The Search Engine Program

The search engine software or program is the final part. When a person requests a search on a keyword or phrase, the search engine software searches the index for relevant information. The software then provides a report back to the searcher with the most relevant web pages listed first. The algorithm-based processes used to determine ranking of results are discussed in greater detail later.

2.2 Human powered directories

These directories compile listings of websites into specific industry and subject categories and they usually carry a short description about the website. Inclusion in directories is a human task and requires submission to the directory producers. Visitors and researchers over the net quite often use these directories to locate relevant sites and information sources. Thus directories assist in structured search. Another important reason is that crawler engines quite often find websites to crawl through their listing and

links in directories. Yahoo and The Open Directory are amongst the largest and most well known directories. LookSmart is a directory that provides results to partner sites such as MSN Search, Excite and others. Lycos is an example of a site that pioneered the search engine but shifted to the Directory model depending on AlltheWeb.com for its listings.

2.3 Hybrid Search Engines

Hybrid Search Engines are both crawler based as well as human powered. In plain words, these search engines have two sets of listings based on both the mechanisms mentioned above. The best example of hybrid search engines is Yahoo, which has got a human powered directory as well as a Search toolbar administered by Google. Although, such engines provide both listings they are generally dominated by one of the two mechanisms. Yahoo is known more for its directory rather than crawler based search engine.

2.4 How do Search Engines rank web pages?

Search engines rank web pages according to the software's understanding of the web page's relevancy to the term being searched. To determine relevancy, each search engine follows its own group of rules. The most important rules are

- The location of keywords on your web page; and

- How often those keywords appear on the page (the frequency)

For example, if the keyword appears in the title of the page, then it would be considered to be far more relevant than the keyword appearing in the text at the bottom of the page.

Search engines consider keywords to be more relevant if they appear sooner on the page (like in the headline) rather than later. The idea is that you'll be putting the most important words – the ones that really have the relevant information – on the page first.

Search engines also consider the frequency with which keywords appear. The frequency is usually determined by how often the keywords are used out of all the words on a page. If the keyword is used 4 times out of 100 words, the frequency would be 4%.

Of course, you can now develop the perfect relevant page with one keyword at 100% frequency - just put a single word on the page and make it the title of the page as well. Unfortunately, the search engines don't make things that simple.

While all search engines do follow the same basic rules of relevancy, location and frequency, each search engine has its own special way of determining rankings. To make things more interesting, the search engines change the rules from time to time so that the rankings change even if the web pages have remained the same.

One method of determining relevancy used by some search engines (like HotBot and Infoseek), but not others (like Lycos), is the Meta tags. Meta tags are hidden HTML

codes that provide the search engine spiders with potentially important information like the page description and the page keywords.

Meta tags are often labeled as the secret to getting high rankings, but Meta tags alone will not get you a top 10 ranking. On the other hand, they certainly don't hurt. Detailed information on meta-tags and other ways of improving search engine ranking is given later in this chapter.

In the early days of the web, webmasters would repeat a keyword hundreds of times in the Meta tags and then add it hundreds of times to the text on the web page by making it the same color as the background. However, now, major search engines have algorithms that may exclude a page from ranking if it has resorted to "keyword spamming"; in fact some search engines will downgrade ranking in such cases and penalize the page.

Link analysis and 'clickthrough' measurement are certain other factors that are "off the page" and yet crucial in the ranking mechanism adopted by some leading search engines. This is quickly emerging as the most important determinant of ranking, but before we study this, we must first look at the most popular search engines and then look at the various steps you can take to improve your success at each of the stages – spidering, indexing and ranking.

2.5 Most popular Search Engines

2.5.1 The most popular engines: Yahoo, Google, and Inktomi

Google is a privately held company that was founded by two Stanford graduates, Larry Page and Sergey Brin in 1998. Dr. Eric Schmidt, the CEO joined in 2001 and by the end of the year the company had shown a profit.

Yahoo has shown sales of around \$ 225 million in the second quarter this year (inching close to the 1 billion mark over the year), and a net income of \$ 21.4 million. Yahoo is a portal and not a pure play search engine company. It was founded by Jerry Yang.

Google is the search engine that powers the search directory for Yahoo. This partnership started in the year 2000 and recently there was a report that the contract is being extended. Last year, Yahoo paid Google about \$7.2 million for Web search services. Inktomi has been a contender too for Yahoo's business. Google also provides an Apple-specific Search Engine specifically tailored to deliver highly targeted results related to Apple Computer and the Macintosh computing platform.

The Apple-specific search engine, located at www.google.com/mac.html, makes searching for everything from Apple's corporate information to product-related news faster and easier.

Inktomi has a robust networking business and a foothold in enterprise search. However, it recently posted deep losses. The company reported a wider net loss in the second quarter

2002, with lower revenue. Its loss broadened to \$104 million or 72 cents a share, from \$58.3 million, or 46 cents a share, a year earlier. Revenue fell to \$30.8 million from \$39.5 million a year earlier.

To stay healthy and competitive in consumer search, Inktomi introduced in the last year a program that generates fees from Web sites listed in its database. Inktomi charges companies such as Amazon.com and eBay to list more than 1,000 Web addresses; they might pay anywhere from 5 cents to 40 cents per click when Web surfers jump to their pages from Inktomi's database. The revenue generated from paid inclusion is shared with partners such as MSN and Overture.

For March 2003, according to a study by Jupiter Media Metrix, there were an estimated 114 million Internet users online in the US at work or at home, 80 percent of whom are estimated to have made some type of search request during the month.

2.5.2 AltaVista

AltaVista is one of the oldest and most well-known search engines. It was launched in December 1995. It was owned by Digital, then run by Compaq (which purchased Digital in 1998), then spun off into a separate company, which is now controlled by CMGI. Over the years it has lost its prominent position to Yahoo and Google.

George McMillan is the CEO of CMGI, Inc. that owns AltaVista. According to the last financial report (fiscal 2001) by parent company CMGI, the "search and portals" division of the company, of which the majority comprises AltaVista's operations, generated \$37 million in revenue with an operating loss of \$72.7 million.

In March 2002 AltaVista has launched its new release of Enterprise Search v2.0 software that it sells to the Enterprise search market, similar to Inktomi and Verity. AltaVista recently launched important freshness and relevancy initiatives, crawling key areas of the Internet four times per day and increasing relevancy by 40%.

AltaVista also executed an agreement pursuant to which Overture Services, Inc. will distribute its 'Pay-For-Performance' search listings on AltaVista's North American sites. The companies signed similar agreements in both Germany and the U.K. during the third quarter. Under the terms of these agreements, AltaVista and Overture will share the revenue resulting from user click-throughs to Overture's customers' Web sites.

2.5.3 Overture (formerly GoTo)

Overture returns search results based on the amount of money an advertiser has paid. This system has made Overture one of the few profitable Net businesses. Yahoo has already signed a three-year deal with Overture to provide paid search results. As this write up does not cover pay per click advertising, we are not focusing much on Overture.

2.5.4 AOL Search

America Online signed a multiyear pact with Google for Web search results and accompanying ad-sponsored links, ending relationships with pay-for-performance service Overture Services and Inktomi, its algorithmic search provider of nearly three years

2.5.5 Verity

We thought it worth mentioning Verity in this, although this is not relevant to search engine optimization. Verity is amongst the leading providers of enterprise search software (Inktomi is another well known player in this). This software is used by numerous enterprises for offering information search within their own sites, portals, intranets/extranets and e-commerce sites. Several software OEMs providing portal or enterprise software also bundle such software in their offerings. Verity (VRTY) went public in 1995 and has achieved sales worth \$ 145 million in fiscal 2001. It is a profitable company. If you wish to provide search facilities to visitors within your site, this may be an option, particularly if you have a large site with lots of information.

2.5.6 Meta Search Engines

Dogpile

Dogpile is a meta-search engine that searches the Internet's top search engines such as About, Ask Jeeves, FAST, FindWhat, Google, LookSmart, Overture and many more. With one single, powerful search engine, you get more relevant and comprehensive results. When you use Dogpile, you are actually searching many search engines simultaneously.

Dogpile, founded in 1996, is the most popular Internet meta-search engine. The site joined the InfoSpace Network in 2000 and is owned and operated by InfoSpace, Inc.

Mamma

Mamma.com is a "smart" meta-search engine - every time you type in a query Mamma simultaneously searches a variety of engines, directories, and deep content sites, properly formats the words and syntax for each, compiles their results in a virtual database, eliminates duplicates, and displays them in a uniform manner according to relevance. It's like using multiple search engines, all at the same time.

Created in 1996 as a master's thesis, Mamma.com helped to introduce meta-search to the Internet as one of the first of its kind. Due to its quality results, and the benefits of meta-search, Mamma grew rapidly through word of mouth, and quickly became an established presence on the Internet. Mamma.com's ability to gather the best results available from top Internet sources and to procure an impressive array of advertisers during the Internet boom of the late 1990s caught the interest of many potential investors. In late 1999

Intasys Corporation (Nasdaq: INTA) invested in a 69% stake of the company and in July 2001 bought the remaining shares and now fully owns Mamma.com.

IxQuick

A search engine is only a tool. It's only as smart as the people who master-minded it. If **IxQuick™** is smarter and more relevant in its findings than other metasearch engines (and many highly qualified people would agree that it is), it's because it was conceived to meet the needs of some of the most inquisitive minds in the world - those of researchers, educators, scholars and the world scientific community.

When a user enters a query at the **IxQuick.com** website, their powerful proprietary technology simultaneously queries ten major search engines and properly formats the words and syntax for each source being probed.

IxQuick™ then creates a virtual database, organizes the results into a uniform format and presents them by relevance and source. In this manner, **IxQuick.com** provides users with highly relevant and comprehensive search results.

3. How major Search Engines work – crawling policies and algorithms

3.1 How Google Works

Google has a comprehensive and highly developed technology, a straightforward interface and a wide-ranging array of search tools which enable the users to easily access a variety of information online. Google users can browse the web and find information in various languages, retrieve maps, stock quotes and read news, search for a long lost friend using the phonebook listings available on Google for all of US cities and basically surf the 3 billion odd web pages on the internet! Google boasts of having world's largest archive of Usenet messages, dating all the way back to 1981. Google's technology can be accessed from any conventional desktop PC as well as from various wireless platforms such as WAP and i-mode phones, handheld devices and other such Internet equipped gadgets.

The web search technology offered by Google is often the technology of choice of the world's leading portals and websites. It has also benefited the advertisers with its unique advertising program that does not hamper the web surfing experience of its users but still brings revenues to the advertisers.

Google's Web Search Technology

When you search for a particular keyword or a phrase, most of the search engines return a list of page in order of the number of times the keyword or phrase appears on the website. Google web search technology involves the use of its indigenously designed PageRank Technology and hypertext-matching analysis which makes several instantaneous calculations undertaken without any human intervention. Google's structural design also expands simultaneously as the internet expands.

PageRank technology

PageRank technology involves the use of an equation which comprises of millions of variables and terms and determines a factual measurement of the significance of web pages and is calculated by solving an equation of 500 million variables and more than 3 billion terms. Unlike some other search engines, Google does not calculate links but utilizes the extensive link structure of the web as an organizational tool. When the link to a Page, lets say Page B is clicked from a Page A, then that click is attributed as a vote towards Page B on behalf of Page A.

Quintessentially, Google calculates the importance of a page by the number of such ‘votes’ it receives. Not only that, Google also assesses the importance of the pages that are involved in the voting process. Consequently, pages that are themselves ahead in ranking and are important in that way also help to make other pages important. One thing to note here is that Google’s technology does not involve human intervention in anyway and uses the inherent intelligence of the internet and its resources to determine the ranking and importance of any page.

Hypertext-Matching Analysis: Unlike its conventional counterparts, Google is a search engine which is hypertext-based. This means that it analyzes all the content on each web page and factors in fonts, subdivisions, and the exact positions of all terms on the page. Not only that, Google also evaluates the content of its nearest web pages. This policy of not disregarding any subject matter pays off in the end and enables Google to return results that are closest to user queries.

Query-handling – The Google Way

Google has a very simple 3-step procedure in handling a query submitted in its search box.

1. When the query is submitted and the enter key is pressed, the web server sends the query to the index servers. Index server is exactly what its name

suggests; it consists of an index much like the index of a book which displays where is the particular page containing the queried term is located in the entire book.

2. After this, the query proceeds to the doc servers, and these servers actually retrieve the stored documents. Page descriptions or “snippets” are then generated to suitably describe each search result.
3. These results are then returned to the user in less than a second!

The Google Dance

Approximately once a month, Google update their index by recalculating the Pageranks of each of the web pages that they have crawled. The period during the update is known as the Google dance.

Because of the nature of PageRank, the calculations need to be performed about 40 times and, because the index is so large, the calculations take several days to complete. During this period, the search results fluctuate; sometimes minute-by-minute. It is because of these fluctuations that the term, **Google Dance**, was coined. The dance usually takes place sometime during the last third of each month.

Google has two other servers that can be used for searching. The search results on them also change during the monthly update and they are part of the Google dance.

For the rest of the month, fluctuations sometimes occur in the search results, but they should not be confused with the actual dance. They are due to Google's fresh crawl and to what is known "Everflux".

Google has two other searchable servers apart from www.google.com. They are www2.google.com and www3.google.com. Most of the time, the results on all 3 servers are the same, but during the dance, they are different.

For most of the dance, the rankings that can be seen on [www2](http://www2.google.com) and [www3](http://www3.google.com) are the new rankings that will transfer to [www](http://www.google.com) when the dance is over. Even though the calculations are done about 40 times, the final rankings can be seen from very early on. This is because, during the first few iterations, the calculated figures merge to being close to their final figures.

You can see this with the Pagerank Calculator by checking the Data box and performing some calculations. After the first few iterations the search results on [www2](http://www2.google.com) and [www3](http://www3.google.com) may still change, but only slightly.

During the dance, the results from [www2](http://www2.google.com) and [www3](http://www3.google.com) will sometimes show on the [www](http://www.google.com) server, but only briefly. Also, new results on [www2](http://www2.google.com) and [www3](http://www3.google.com) can disappear for short periods. At the end of the dance, the results on [www](http://www.google.com) will match those on [www2](http://www2.google.com) and [www3](http://www3.google.com).

This Google Dance Tool allows you to check your rankings on www, www2 and www3 and on all 9 datacenters simultaneously.

How to maximize Google's Search Features?

Find It ASAP

The Google Web Directory works in combination of the Google Search Technology and the Netscape Open Directory Project which makes it possible to search the Internet organized by topic. Google displays the pages in order of the rank given to it using the PageRank Technology. It not only searches the titles and descriptions of the websites, but searches the entire content of sites within a category, which ultimately delivers a comprehensive search to the users. Google also has a fully functional web directory which categorizes all the searches in order.

In case you are feeling lucky

The I'm Feeling Lucky™ search button is recommended when searching for a highest ranked web page for a particular search. This saves time in searching for a webpage.

The multi-faceted Google Toolbar

The Google Toolbar™ can seamlessly integrate with a user's web browser and be of quick assistance.

Pandora's Box

- Google enables its users to search for U.S. street maps immediately by just typing the street name in the query box.
- Latest stock quotes are just a click away. Just type in the company ticker symbol or the name of one of the stock indices stock and mutual fund information and Google will return the relevant information in association with high-profile financial and trading concerns.

Cache me if you can

Google takes a snapshot of each page it examined as it crawls the web and stores or caches them as a back-up in case the original page is unavailable. This cached link always displays the page in the same manner as it was indexed and this is used by Google to match the relevancy of the page to the query submitted by the user. The "Cached" link will be missing for sites that have not been indexed, as well as for sites whose owners have requested Google not to cache their content.

Google Web Search Features

Google offers a variety of special features which helps users to find exactly what they are looking which is all in addition to providing easy access to more than 3 billion web pages. The following is an overview of its key features:

- **Calculator**

Google has a built-in calculator function which can be used to calculate mathematical expressions involving basic arithmetic, more complicated math, units of measure and conversions, physical constants and even hexadecimal and binary numbering systems. You can simply enter the expression you'd like evaluated in the search box and hit the Enter key or click the Google Search button.

- **Dictionary Definitions**

When searching about any particular term, if the Google database has a definition or meaning for the term, then it will be highlighted with an underline on the results page. This definition is derived in association with a reliable dictionary source.

- **File Types**

In addition to HTML files, Google search also supports 12 other formats such as PDF, Microsoft Office, PostScript, Corel WordPerfect, Lotus 1-2-3, and others. Additionally, Google also offers the user the ability to "View as HTML", which allows users to view these files in case the corresponding software is not installed on the user's PC. It also eliminates the hazards of opening a virus-infected document.

- **News Headlines**

While searching for a particular term, if that term is also in any of the current news, it is displayed as a separate news link on the results page. This is derived from various news providers who work in association with Google and who let Google monitor them.

A new offering – Advanced News Search

A new offering from Google - Advanced News Search, allows visitors to scour headlines by date, location, exact phrases or publication. People can use it retrieve articles from more than 4,500 news outlets publishing on the Web. Advanced News Search lets visitors search for headlines using several parameters. Among other features, people can locate stories that contain an exact phrase, within the United States or abroad, or written by a specific publisher.

- **Similar Pages**

The results page also displays a link for ‘similar pages’ which uses the GoogleScout technology to explore the web for similar pages. This is particularly helpful if you have hit upon a page which has relevant content, but you want something similar but more.

- **Web Page Translation**

This feature is particularly helpful if your search has non-English results. Google offers a facility to automatically translate a page for you in English. Currently, Google supports Italian, French, Spanish, German, and Portuguese languages.

- **SafeSearch Filtering**

Google provides a SafeSearch option to filter pornographic contents from its results page. This is especially useful for shared computers which need to be protected for children surfing the Internet. Google’s technology tries to check keywords and phrases, URLs and Open Directory categories and eliminate these from the search results.

Submitting your URL to Google

Google is primarily a fully-automatic search engine with no human-intervention involved in the search process. It utilizes robots known as 'spiders' to crawl the web on a regular basis for new updates and new websites to be included in the Google Index. This robot software follows hyperlinks from site to site. Google does not require that you should submit your URL to its database for inclusion in the index, as it is anyway done automatically by the 'spiders'. However, manual submission of URL can be done by going to the Google website and clicking the related link. One important thing here is that Google does not accept payment of any sort for site submission or improving page rank of your website. Also, submitting your site through the Google website does not guarantee listing in the index.

Cloaking

Sometimes, a webmaster might program the server in such a way that it returns different content to Google than it returns to regular users, which is often done to misrepresent search engine rankings. This process is referred to as cloaking as it conceals the actual website and returns distorted webpages to search engines crawling the site. This can mislead users about what they'll find when they click on a search result. Google highly disapproves of any such practice and might place a ban on the website which is found guilty of cloaking.

Google Guidelines

Here are some of the important tips and tricks that can be employed while dealing with Google.

Do's

- A website should have crystal clear hierarchy and links and should preferably be easy to navigate.
- A site map is required to help the users go around your site and in case the site map has more than 100 links, then it is advisable to break it into several pages to avoid clutter.
- Come up with essential and precise keywords and make sure that your website features relevant and informative content.
- The Google crawler will not recognize text hidden in the images, so when describing important names, keywords or links; stick with plain text.
- The TITLE and ALT tags should be descriptive and accurate and the website should have no broken links or incorrect HTML.
- Dynamic pages (the URL consisting of a '?' character) should be kept to a minimum as not every search engine spider is able to crawl them.

- The robots.txt file on your web server should be current and should not block the Googlebot crawler. This file tells crawlers which directories can or cannot be crawled.

Don'ts

- When making a site, do not cheat your users, i.e. those people who will surf your website. Do not provide them with irrelevant content or present them with any fraudulent schemes.
- Avoid tricks or link schemes designed to increase your site's ranking.
- Do not employ hidden texts or hidden links.
- Google frowns upon websites using cloaking technique. Hence, it is advisable to avoid that.
- Automated queries should not be sent to Google.
- Avoid stuffing pages with irrelevant words and content. Also don't create multiple pages, subdomains, or domains with significantly duplicate content.
- Avoid "doorway" pages created just for search engines or other "cookie cutter" approaches such as affiliate programs with hardly any original content.

Google Services

The following are some of the popular and innovative services offered by Google and there are always improvisations in it.

Google Answers

Google Answers is an interesting cross between 'online marketplace' and probably a 'virtual classroom'. Those who wish to participate must register with Google Answers. Here, the researchers who have considerable expertise in online researching provide answers to the queries posted by other users for a fee. When a user posts a question, he or she also needs to mention the price the user is willing to pay in case the question is answered. When the question is answered by any user, then the payment is made accordingly to the user answering the question. Moreover, the questions and the discussion that ensues will be publicly viewable and other registered users can also share their opinions and insights.

There is a non-refundable listing fee of \$0.50 per question plus an additional 'price' you set for your question that reflects how much you're willing to pay for an answer. Three-quarters of your question price goes directly to the Researcher who answers your question; the remaining 25 percent goes to Google to support the service.

Google Groups

Google Groups is an online discussion forum and it contains the entire archive of Usenet discussion groups dating back to 1981. These discussions cover the full range of human

dissertation and present a fascinating look at evolving viewpoints, debate and advice on every subject from politics to technology. Users can access all of this information all in a database that contains more than 800 million posts by using the search feature of Google.

Google's Image Search

Google offers a wide collection of images from around the web; its comprehensive database consists of more than 425 million images. All a user has to do is to enter a query in the image search box, then click on the "Search" button. On the results page, by clicking the thumbnail a larger version of the image can be seen, as well as the web page on which the image is located. By default, Google's Image Search uses its mature content filter on the initial search by any user. The filter removes many adult images but it cannot guarantee that all such content will be filtered out. It is not possible to ensure with 100% accuracy that all mature content will be removed from image search results using filters.

Google analyzes the text on the page near the image, the image caption and dozens of other factors which enables it to determine the image content. Google also utilizes several sophisticated algorithms which make it possible to remove duplicates and it in turn ensures that the highest quality images are presented first in the results. Google's Image search supports all the complex search strategies like Boolean operators, etc.

Google's Catalog Search

Google offers a unique service in the form of its Catalog Search. Google's Catalog Search has made it easy to find information published in mail-order catalogs that were not previously available online. It includes the full content of hundreds of mail-order catalogs selling everything from industrial adhesives to clothing and home furnishings. Google's Catalog Search can help you if you are looking to buy for either yourself or for your business.

The printed copies of catalogs are scanned and the text portion is converted into a format which makes it easy for users to search for the catalog. The same sophisticated algorithm employed by the Google Web Search is then employed to search for catalogs. This makes sure that most recent and relevant catalogs are displayed. Google is not associated with any catalog vendors and is not liable for any misuse of this service on part of the users.

Froogle

The word 'froogle' is a combination of the word 'frugal' which means 'penny-wise' or 'economical' and of course 'Google'. Currently in its beta version, or testing format, Froogle is a recent concept put forth by Google. Google's spidering software crawls the

web looking for information about products for sale online. It does so by focusing entirely on product search and applying the power of Google's search technology to locate stores that sell items you want and consequently pointing you to that specific store.

Just like the Google Web Search, Froogle also ranks store sites based only on their relevance to the search terms entered by the users. Google does not accept payment for placement within their actual search results. Froogle also includes product information submitted electronically by merchants. Its search results are automatically generated by Google's ranking software.

3.2 How AltaVista Works

AltaVista has an index that is built by sending out a crawler (a robot program) that captures text and brings it back.

The main crawler is called "Scooter." Scooter sends out thousands of threads simultaneously. 24 hours a day, 7 days a week, Scooter and its cousins access thousands of pages at a time, like thousands of blind users grabbing text, pulling it back, throwing it into the indexing machines so the next day that text can be in the index. And at the same time, they pull off, from all those pages, every hyperlink that they find, to put in a list of where to go to next.

In a typical day Scooter and its cousins visit over 10 million pages. If there are a lot of hyperlinks from other pages to yours, that increases your chances of being found. But if this is your own personal site, or if this is a brand new Web page, that's not too likely.

AltaVista has an incredibly large database of Web sites, such that searches often return hundreds of thousands of Web site matches. AltaVista's spider goes down about three pages into your site. This is important to remember if you have different topical pages that won't be found within three clicks of the main page. You will have to index them separately.

You cannot tell Alta Vista how to index your site, it is all done via their spider, but you can go to their site and give the spider a nudge by submitting specific pages. That way, AltaVista's spider knows to visit that page and index it. Once you have done that, it's all up to your META tags and your page's content! AltaVista's spider may revisit your site each month after its initial visit.

AltaVista ranking algorithms reward keywords in the <TITLE> tag. If a keyword is not in a title tag, it will likely not appear anywhere near the top of the search results!

AltaVista also rewards keywords near one another, and keywords near the beginning of a page

Add a Page

Adding a page through AltaVista's Add URL form doesn't guarantee that the page would be listed. It usually takes around 4 to 6 weeks to show up. You don't have to have any special authority to "add a page." This is not a directory, like Yahoo!, where the information provider has to submit information and has to prove they are who they say they are. You do not have to do this with AltaVista. It will go and check and bring back whatever text it finds at that address.

If you give it a URL for a page that doesn't exist, it will come back with Error 404, which means there is no such page. If that page was in the index, it will remove that page from the index the next day.

This is very important from several perspectives. Say you have changed the directory structure at your Web site. First, you should go to AltaVista and Add a Page for all the old addresses to remove the old information from the index. Then you should Add a Page for all the new addresses. Also, if you made an embarrassing typo or posted a document that you shouldn't have, and removed that page from the Web, you can Add URL for that page at AltaVista to make sure the information is not perpetuated in the index.

What AltaVista doesn't Index

AltaVista doesn't index everything. In fact, features that Web designers may add to sites at great expense may block crawlers, meaning that those pages will never be indexed and

never be found through search engines. As a result, those sites may end up spending far more on promotion than they would have had to otherwise.

Here are some pages AltaVista doesn't index. This only highlights the importance of using plain text for your web pages.

First, sites that require any kind of registration or password lock out AltaVista. Keep in mind that a web crawler cannot fill out a form of any kind. If you need to fill out a form to get to the next page, the crawler halts right there. If you would like to gather information about your users/members but would also like your pages to be indexed, make the registration optional.

Similarly, the AltaVista crawler cannot get content from a database, because it cannot fill out a form. If the content of your database is largely text, you might consider creating plain text static HTML pages with that same content, so it can be indexed and found.

Dynamic pages also block AltaVista spiders. While it's great to give visitors to your site unique experiences, tailored to their needs, the techniques you use to do that could stop most search engines including AltaVista from indexing your content and hence could greatly reduce your potential traffic. Dynamically generated pages are created on the fly from a variety of elements held in databases. When the AltaVista crawler arrives at such a page, it captures the content but halts immediately, and will not follow the links,

because it sees ahead of it an infinite number of pages ahead -- a black hole that would bring it to a crash.

Active Server Pages (.asp) with question marks in their URLs (indicating that the page is a script for the construction of a page, rather than just static content) fall into this category.

If you have information inside frames, that will probably prove to be a hindrance, but is not an absolute barrier. AltaVista indexes the outside of the frame as a distinct page. It will also index each pane of the frame window as a separate page. That means that if the content matching a query is in a pane, when visitors clicking on those links will see the pane and only the pane -- not the full page as it was designed. So if you want visitors from search engines to experience your pages the way they were intended to be seen, you should have non-frames as well as frames versions of those pages; and submit the non-frames versions with Add URL.

AltaVista also can't index text that is embedded in graphics. Search engines simply cannot "see" the text unless the Webmaster put ALT text behind the picture, describing it and listing those important words. But pictures, as pictures, can be indexed for Image search at AltaVista.

Text that appears in multi-media files (audio and video) cannot be indexed. But those same files can be indexed at AltaVista for Multimedia search.

Information that is generated by Java applets or in XML coding cannot be indexed.

Acrobat files cannot be indexed either. But technology exists that will enable AltaVista to convert those files to indexable form.

Exceptionally large pages also present a problem at AltaVista. As a pragmatic compromise, intended to help optimize the performance of AltaVista, they fully index the first 64 Kbytes of text on any single page. They will harvest the hyperlinks from the whole document for following up later, but they will only index the first 64 Kbytes. So if you want to post an entire book, it's best to break it up into chapters, and then all the text can be indexed.

Comments, such as <!--change this every Friday-->, aren't indexed at all. Those are intended as private communications, not viewable by Web site visitors, except by using View/Page Source.

Also, consider technical factors. If a site has a slow connection, it might time-out for the crawler. Very complex pages, too, may time out before the crawler can harvest the text.

If you have a hierarchy of directories at your site, put the most important information high, not deep. AltaVista will presume that the higher you placed the information, the

more important it is. And crawlers may not venture deeper than three or four or five directory levels.

Above all remember the obvious - full-text search engines such as AltaVista index text. You may well be tempted to use fancy and expensive design techniques that either block search engine crawlers or leave your pages with very little plain text that can be indexed.

Ranking Rules

The simple rule of thumb is that content counts, and that content near the top of a page counts for more than content at the end. In particular, the HTML title and the first couple lines of text are the most important part of your pages. If the words and phrases that match a query happen to appear in the HTML title or first couple lines of text of one of your pages, chances are very good that that page will appear high in the list of search results.

AltaVista bases its ranking on both static factors (a computation of the value of page independent of any particular query) and query-dependent factors.

It values:

- Long pages, which are rich in meaningful text (not randomly generated letters and words).

- Pages that serve as good hubs, with lots of links to pages that have related content (topic similarity, rather than random meaningless links, such as those generated by link exchange programs or intended to generate a false impression of "popularity").
- The connectivity of pages, including not just how many links there are to a page but where the links come from: the number of distinct domains and the "quality" ranking of those particular sites. This is calculated for the site and also for individual pages. A site or a page is "good" if many pages at many different sites point to it, and especially if many "good" sites point to it.
- The level of the directory in which the page is found. Higher is considered more important. If a page is buried too deep, and the crawler simply won't go that far and will never find it.

These static factors are recomputed about once a week, and new good pages slowly percolate upward in the rankings. Note that there are advantages to having a simple address and sticking to it, so others can build links to it, and so you know that it's in the index

Query-dependent factors include:

- The HTML title.
- The first lines of text.
- Query words and phrases appearing early in a page rather than late.
- Meta tags, which are treated as ordinary words in the text, but like words that appear early in the text (unless the meta tags are patently unrelated to the content on the page itself, in which case the page will be penalized)
- Words mentioned in the "anchor" text associated with hyperlinks to your pages. (E.g., if lots of good sites link to your site with anchor text "breast cancer" and the query is "breast cancer," chances are good that you will appear high in the list of matches.)

AltaVista's policy on doorway pages and cloaking

AltaVista is opposed to doorway pages and cloaking. It considers doorway and cloaked pages to be spam and encourages people to use other avenues to increase the relevancy of their pages. A description of doorway pages and cloaking is given later on in this guide.

Meta tags

Though it indexes Meta tags, considering them to be regular text, AltaVista claims it doesn't give them priority over HTML titles and other text. Though you should use meta tags in all your pages, some webmasters claim their doorway pages for AltaVista rank better when they don't use them.

If you do use Meta tags, make your description tag no more than 150 characters and your keywords tag no more than 1,024 characters long.

Keywords in the URL and file names

It's generally believed that AltaVista gives some weight to keywords in filenames and URL names. If you're creating a file, try to name it with keywords.

Keywords in the ALT tags

AltaVista indexes ALT tags, so if you use images on your site, make sure to add them. ALT tags should contain more than the image's description. They should include keywords, especially if the image is at the top of the page. ALT tags are explained later.

Page Length

There's been some debate about how long doorway pages for AltaVista should be. Some webmasters say short pages rank higher, while others argue that long pages are the way to go. According to AltaVista's help section, it prefers long and informative pages. We've found that pages with 600-900 words are most likely to rank well.

Frame support

AltaVista has the ability to index frames, but it sometimes indexes and links to pages intended only as navigation. To keep this from happening to you, submit a frame-free site map containing the pages that you want indexed. You may also want to include a "robots.txt" file to prohibit AltaVista from indexing certain pages.

AltaVista's Search Features

AltaVista offers a wide range of search features. Most of these options are available in its "Advanced Search" section.

- Boolean search - Full Boolean Search support. AND (+), OR, AND NOT (-) (instead of NOT) Search terms can be nested.
- Phrase - Available. Put quotes around the phrase, such as "New York Times"
- Proximity - Available. NEAR operator means within ten words of one another. Can be nested with other tags.
- Word Stemming - Available. You can use the wild card (*) at the end or in the middle of a word.

- Capitalization - If you search in upper case, AltaVista searches in upper case only. Lower case words and phrases search for upper and lower case, and will therefore yield more results.

- Field Search - The following options are available:
 - Applet: searches for the name of an applet
 - Domain: specifies the domain extension, such as .com
 - Host: searches for pages within a particular site
 - Image: searches for an image name
 - Link: searches for pages that link to the specified site
 - Object: search engines - searches for the name of an object
 - Text: excludes Meta tags information
 - Title: search in the HTML title only
 - URL: searches for sites that have a specified word in the URL

- Date Searching - Available under Advanced Search section.

- Search within results - Available. This option is offered after each search.

- Media Type searching - Available for Images, Music/MP3, and Video.

- Language Searching - AltaVista has very extensive language support. It supports around 30 languages.

3.3 How Teoma Works

Teoma adds a new dimension and level of authority to search results through its approach, known as **Subject-Specific PopularitySM**.

Instead of ranking results based upon the sites with the most links leading to them, Teoma analyzes the Web as it is organically organized—in naturally-occurring communities that are about or related to the same subject—to determine which sites are most relevant. Teoma's search technology can locate communities on the Web within their specific subject areas, as they actually exist.

To determine the authority—and thus the overall quality and relevance—of a site's content, Teoma uses **Subject-Specific PopularitySM**. Subject-Specific Popularity ranks a site based on the number of same-subject pages that reference it, not just general popularity. In a recent test performed by *Search Engine Watch*, Teoma's relevance grade was raised to an "A" following the integration of Teoma 2.0.

Teoma 2.0: Evolution and Growth

In early 2003, Teoma 2.0 was launched. The enhanced version represents a major evolution in terms of improvements to relevance and an expansion of the overall advanced search functionalities. Below are detailed explanations for the improvements made in this version:

More Communities

Like social networks in the real world, the Web is clustered into local communities. Communities are groups of Web pages that are about or are closely related to the same subject. Teoma is the only search technology that can view these communities as they naturally occur on the Web (displayed under the heading "Refine" on Teoma.com). This method allows Teoma to generate more finely tuned search results. In other words, Teoma's community-based approach reveals a 3-D image of the Web, providing it with more information about a particular Web page than other search engines, which have only a one-dimensional view of the Web.

Web-Based Spell Check

Teoma's proprietary Spell Check technology identifies query misspellings and offers corrections that help improve the relevance and precision of search results. The Spell Check technology, developed by Teoma's team of scientists, leverages the real-time content of the Web to determine the correct spelling of a word.

Dynamic DescriptionsSM

Dynamic Descriptions enhance search results by showing the context of search terms as they actually appear on referring Web pages. This feature provides searchers with information that helps them to determine the relevance of a given Web page in association with their query.

Advanced Search Tools

Teoma's Advanced Search tools allow searchers to search using specific criteria, such as exact phrase, page location, geographic region, domain and site, date, and other word filters. Users can also search using 10 Western languages, including Danish, Dutch, English, French, German, Italian, Norwegian, Portuguese, Spanish and Swedish. A link to Teoma's Advanced Search tools can be found next to the search box on Teoma.com.

The Teoma Algorithm

In addition to utilizing existing search techniques, Teoma applies what they call authority, a new measure of relevance, to deliver search results. For this purpose, Teoma employs three proprietary techniques:

Refine, Results and Resources.

Refine

First, Teoma organizes sites into naturally occurring communities that are about the subject of each search query. These communities are presented under the heading "**Refine**" on the Teoma.com results page. This tool allows a user to further focus his or her specific search.

For example, a search for "Soprano" would present a user with a set of refinement suggestions such as "Marie-Adele McArther" (a renowned soprano), "Three Sopranos" (the operatic trio), "The Sopranos" (the wildly-popular HBO television show) as well as several other choices. No other technology can dynamically cluster search results into the actual communities as they exist on the Web.

Results

Next, after identifying these communities, Teoma employs a technique called **Subject-Specific PopularitySM**. Subject-Specific Popularity analyzes the relationship of sites within a community, ranking a site based on the number of same-subject pages that reference it, among hundreds of other criteria. In other words, Teoma determines the best answer for a search by asking experts within a specific subject community about who they believe is the best resource for that subject. By assessing the opinions of a site's

peers, Teoma establishes authority for the search result. Relevant search results ranked by Subject-Specific Popularity are presented under the heading "**Results**" on the Teoma.com results page.

In some instances companies pay to have their Web sites included within Teoma's dataset, otherwise known as the Teoma Index. Like all Web sites, these sites are processed through Teoma's search algorithms and are not guaranteed placement in the results. This ensures that relevancy is the primary driver of results.

Resources

Finally, by dividing the Web into local subject communities, Teoma is able to find and identify expert resources about a particular subject. These sites feature lists of other authoritative sites and links relating to the search topic.

For example, a professor of Middle Eastern history may have created a page devoted to his collection of sites that explain the geography and topography of the Persian Gulf. This site would appear under the heading "**Resources**" in response to a Persian Gulf-related query. No previous search technology has been able to find and rank these sites.

Sponsored Links

Search results appearing under the heading "**Sponsored Links**" are provided by Google[®], a third party provider of pay for performance search listings. Google generates highly relevant sponsored results by allowing advertisers to bid for placement in this area based on relevant keywords. These results, which are powered by Google's advanced algorithms, are then distributed across the Internet to some of the world's most popular and well-known Web sites, including Teoma.com and Ask Jeeves.

Other factors

Boolean Searching

Limited Boolean searching is available. Teoma defaults to an AND between search terms and supports the use of - for NOT. Either OR or ORR can be used for an OR operation, but the operator must be in all upper case. Unfortunately, no nesting is available.

Proximity Searching

Phrase searching is available by using "double quotes" around a phrase or by checking the "Phrase Match" box. Teoma also supports phrase searching when a dash is used between words with no spaces. Until Nov. 2002, Teoma's help page stated that "Teoma returns results which exactly or closely matches the given phrase" which meant that not

all phrases matches will necessarily be accurate. As of Nov. 2002, that appears to have been corrected and phrase searching now works properly.

Truncation

No truncation is currently available.

Case Sensitivity

Searches are not case sensitive. Search terms entered in lowercase, uppercase, or mixed case all get the same number of hits.

Stop Words

Teoma does ignore frequently-occurring words such as 'the,' 'of', 'and', and 'or'. However, like at Google, these stop words can be searched by putting a + in front of them or by including them within a phrase search.

Sorting

By defaults, sites are sorted in order of perceived relevance. They also have site collapsing (showing only two pages per site with the rest link via a “More Results” message. There is no option for sorting alphabetically, by site, or by date.

Display

Teoma displays the title (roughly first 60 characters), a two line keyword-in-context extract from the page, and the beginning of the URL for each hit. Some will also have a link to "Related Pages" which finds related records based on identifying Web communities by analyzing link patterns. Two other sections displayed are the "Refine" section (formerly folders) that suggest other related searches based on words that Teoma uses to identify communities on the Web and the "Resources: Link collections from experts and enthusiasts" (formerly "Experts' Links") which are Web pages that include numerous links to external resources -- metasites or Internet resource guides. Some "Sponsored Links" may show up at the top. These are ads from the Google AdWords program.

Teoma will only display 10 Web page records at a time; however, up to a 100 at a time can be displayed through a change in the preferences and on the advanced search page. Teoma may also display up to 10 metasites under the "Resources" heading and up to 6 Refine suggestions.

3.4 How Inktomi Works

Inktomi is one of the most popular crawler based search engines. Inktomi is a crawler-based search engine. However, it does not make its index available to the public through its own site like other crawler-based search engines, such as Lycos or Alltheweb. Inktomi licenses other companies to use its search index. These companies are then able to provide search services to their visitors without having to build their own index.

It uses a robot named Slurp to crawl and index web pages.

Slurp – The Inktomi Robot

Slurp collects documents from the web to build a searchable index for search services using the Inktomi search engine, including Microsoft and HotBot. Some of the characteristics of Slurp are given below:

Frequency of accesses

Slurp accesses a website once every five seconds on average. Since network delays are involved it is possible over short periods the rate will appear to be slightly higher, but the average frequency generally remains below once per minute.

robots.txt

Slurp obeys the Robot Exclusion Standard. Specifically, Slurp adheres to the 1994 Robots Exclusion Standard (RES). Where the 1996 proposed standard disambiguates the 1994 standard, the proposed standard is followed.

Slurp will obey the first record in the robots.txt file with a User-Agent containing "Slurp". If there is no such record, it will obey the first entry with a User-Agent of "*".

This is discussed in detail later in this book.

NOINDEX meta-tag

Slurp obeys the NOINDEX meta-tag. If you place

```
<META NAME="robots" CONTENT="noindex">
```

in the head of your web document, Slurp will retrieve the document, but it will not index the document or place it in the search engine's database.

Repeat downloads

In general, Slurp would only download one copy of each file from your site during a given crawl. Occasionally the crawler is stopped and restarted, and it re-crawls pages it has recently retrieved. These re-crawls happen infrequently, and should not be any cause for alarm.

Searching the results

Slurp crawls from websites to the Inktomi search engines immediately. The documents are indexed and entered into the search database in quick time.

Following links

Slurp follows HREF links. It does not follow SRC links. This means that Slurp does not retrieve or index individual frames referred to by SRC links.

Dynamic links

Slurp has the ability to crawl dynamic links or dynamically generated documents. It will not, however, crawl them by default. There are a number of good reasons for this. A couple of reasons are that dynamically generated documents can make up infinite URL spaces, and that dynamically generated links and documents can be different for every retrieval so there is no use in indexing them.

Content guidelines for Inktomi

Given here are the content guidelines and policies for Inktomi. In other words, listed below is the content Inktomi indexes and the content it avoids.

Inktomi indexes:

- Original and unique content of genuine value
- Pages designed primarily for humans, with search engine considerations secondary

- Hyperlinks intended to help people find interesting, related content, when applicable
- Metadata (including title and description) that accurately describes the contents of a Web page
- Good Web design in general

Inktomi avoids:

- Pages that harm accuracy, diversity or relevance of search results
- Pages dedicated to directing the user to another page
- Pages that have substantially the same content as other pages
- Sites with numerous, unnecessary virtual hostnames
- Pages in great quantity, automatically generated or of little value
- Pages using methods to artificially inflate search engine ranking
- The use of text that is hidden from the user
- Pages that give the search engine different content than what the end-user sees
- Excessively cross-linking sites to inflate a site's apparent popularity
- Pages built primarily for the search engines
- Misuse of competitor names
- Multiple sites offering the same content
- Pages that use excessive pop-ups, interfering with user navigation
- Pages that seem deceptive, fraudulent or provide a poor user experience

Inktomi's policies are designed to ensure that poor-quality pages do not degrade the user experience in any way. As with Inktomi's other guidelines, Inktomi reserves the right, at its sole discretion, to take any and all action it deems appropriate to insure the quality of its index.

Inktomi encourages Web designers to focus most of their energy on the content of the pages themselves. They like to see truly original text content, intended to be of value to the public. The search engine algorithm is sophisticated and is designed to match the regular text in Web pages to search queries. Therefore, no special treatment needs to be done to the text in the pages.

They do not guarantee that your web page will appear at the top of the search results for any particular keyword.

How does Inktomi rank web pages?

Inktomi search results are ranked based on a combination of how well the page contents match the search query and on how "important" the page is, based on its appearance as a reference in other web pages.

The quality of match to the query terms is not just a simple text string match, but a text analysis that examines the relationships and context of the words in the document. The

query match considers the full text content of the page and the content of the pages that link to it when determining how well the page matches a query.

Here are a few tips that can make sure your page can be found by a focused search on the Internet:

- Think carefully about key terms that your users will search on, and use those terms to construct your page.
- Documents are ranked higher if the matching search terms are in the title. Users are also more likely to click a link if the title matches what they're looking for. Choose terms for the title that match the concept of your document.
- Use a "description" meta-tag and write your description carefully. After a title, users click on a link because the description draws them in. Placing high in search results does little good if the document title and description do not attract interest.
- Use a "keyword" meta-tag to list key words for the document. Use a distinct list of keywords for each page on your site instead of using one broad set of keywords on every page. (Keywords do not have much effect on ranking, but they do have an effect.)

- Keep relevant text and links in HTML. Placing them in graphics or image maps means search engines can't search for the text and the crawler can't follow links to your site's other pages. An HTML site map, with a link from your welcome page, can help make sure all your pages are crawled.
- Use ALT text for graphics. It's good page design to accommodate text browsers or visually impaired visitors, and it helps improve the text content of your page for search purposes.
- Correspond with webmasters and other content providers and build rich linkages between related pages. Note: "Link farms" create links between unrelated pages for no reason except to increase page link counts. Using link farms violates Inktomi content guidelines, and will not improve your page ranking.

Inktomi's Spamming Policies

Sites that violate the Inktomi content guidelines may be removed from the index. These sites are considered as spam. Inktomi considers techniques such as tiny text, invisible text, keyword stuffing, doorway pages, and fake links as spam.

Pages with no unique text or no text at all may drop out of the index or may never be indexed. If you want a page to appear in web search results, be sure that page includes some unique text content to be indexed.

Inktomi, however, does index dynamic pages. For page discovery, Inktomi mostly follows static links, and the avoidance of dynamically generated href links except in directories disallowed by a /robots.txt exclusion rule is recommended.

Spamming includes:

- Embedding deceptive text in the body of web documents.
- Creating metadata that does not accurately describe the content of web documents.
- Fabricating URLs that redirect to other URLs for no legitimate purpose.
- Web documents with intentionally misleading links
- Cloaking/doorway pages that feed Inktomi crawlers content that is not reflective of the actual page
- Creating inbound links for the sole purpose of boosting the popularity score of the URL
- The misuse of third party affiliate or referral programs

Click popularity measurement

As mentioned earlier, Inktomi measures the click popularity of web pages while deciding the rank of a web page. Click popularity is the number of times the surfers click on your web page listing and how long they stay in your site.

The number of click on your site's listing can be improved by utilizing the title and the Meta tags. These two tags not only help you in attaining a high rank in the search engines, but they also can be utilized to write a good marketing text about your site. The text in the title and Meta description tag appears in the hyperlink listings on the search engine results page. If the text is attractive to the net surfers, the chances of getting more clicks is greater.

Another factor which decides the click popularity factor of your web site is the time that the visitors spend in your site. The secret behind retaining visitors in your web site is the content of your site. Informative and useful content relevant to the search terms will help to retain visitors to your site and make them come back again.

Inktomi's Partner sites

Inktomi provides search results to many search sites. The different search portals may also use results from other information sources, so not all of their results come from the Inktomi search database. These search portals also apply different selection or ranking constraints to their search requests, so Inktomi results at different portals may not be the same.

Following are Inktomi's partner sites:

<http://www.about.com/>

<http://www.bbc.co.uk/>

<http://www.bluewin.ch/>

<http://www.blueyonder.co.uk/>

<http://www.espotting.com/>

<http://www.fi/>

<http://www.goo.ne.jp/>

<http://www.hotbot.com>

<http://www.hotbot.co.uk/>

<http://www.looksmart.com>

<http://search.msn.com/>

<http://www.overture.com/>

<http://www.soneraplaza.fi/>

<http://www.tocc.co.jp/search/>

<http://www.wp.pl/>

How You Can Finally Achieve Top Search Engine Rankings!

4. Introduction to Search Engine Submission

A majority of search engines charge a fee for submission. There are also some good search engines that allow you to submit your site for free. There are a few search engines that offer free submission services. Although, listing of your website is not guaranteed through free submission, it is always recommend submitting your website to a search engine with free submission.

4.1 Search Engine Submission – Getting the Basics Right

Free listing is available with some major search engines such as Google, AltaVista, and Webcrawler. As stated earlier, even after submission of your website, a listing is not guaranteed. Generally, it takes up to 2 to 3 months to be listed after submission. A recent research concluded that the audience potential for websites submitted through free search engines is about 39%.

Most search engines charge a nominal fee for every URL submitted. You are more likely to get favorable results by submitting your website to a search engine with paid inclusion. AskJeeves, Inktomi, AOL, and LookSmart are the most popular search engines in this

category. Once you submit your website with these Search Engines, your website is most likely to be listed within a week. The Audience potential for paid search engines is 100%.

Search engines generally list pages of ten to twenty results per page. Most search items will return thousands, if not hundreds of thousands, of results. The key, however, is to get listed amongst the top results if expect to get any traffic through these search engine results.

There are two methods of submitting your URL to Search Engines. One is to use Search Engine Submission services such as “Submit it”, which is a part of MSN Central. The fee for submitting URLs using this service ranges from \$79 to \$299 per year. Another is to submit your URL by submitting it individually to popular Search Engines and thus avoiding the fee charged by submission services. The prevailing counsel is that manual submissions should be made to the top five search engines or so and one could use automatic submission services for the rest.

It is possible to submit your URL to search engines for free. At the same time, you have to keep in mind that there are premium programs offered by some search engines that assure listing or provide better positioning in ranking. Some use of paid listing programs is recommended if you wish to receive serious traffic on your website from the search engine. Paid Listing Programs are explained later in this chapter. Free submission could result in much lower traffic, a low rank and positioning amongst the results returned, and longer period of time before your website actually shows up in the results returned by the

search engine. Also there is no guarantee that your website would be included with these free services. To avoid these issues some search engines offer “paid participation” that guarantees high traffic and ranking for a fixed fee per year.

4.2 The Submission Budget

Formulating a Search Engine submission budget is crucial. It should be such that you have the best possible combination of free submission, paid submission and paid placement programs.

How much would you like to spend on this exercise? If the budget is limited, options such as some of the paid programs, advertisements, expensive directory listings will have to be forsaken and attention given to getting the best results from limited but focused efforts.

The key is to strike a balance between free and paid programs that yield maximum ROI. You should submit your website to all the free search engines such as Google, AltaVista, and WebCrawler, at least one paid search engine such as Inktomi (\$89 for submission of 3 URLs), and Yahoo Directory (\$299 a year). Apart from these, you should think of submitting your website to a couple of “paid participation” or “paid placement” programs such as Google Adwords and Overture, which are discussed later.

4.3 Submitting to Google

Google is probably the most widely used web crawler search engine. One way of letting Google automatically detect and include your web page URL is as discussed in the above paragraph. Submission to Google through its free listing program can be done using the Add URL form available at

<http://www.google.com/addurl.html>

Having mentioned that, there is no guarantee that Google would include a web page submitted to it by either of the free methods. Also, it might take as long as a month before Google lists your web pages.

Google has advertising program called Adwords that offers you a combination of paid listing and positioning of your pages in search results. Paid listings in Google appear above and to the side of its regular results. These are discussed in detail in the section on Paid Placement programs.

Google allows maximum 5 to 10 web pages from a website to be submitted per day.

Other web crawler search engines with both free as well as paid submission services are discussed below. The free submission method for these search engines is very similar to those described in the case of Google.

4.4 Submitting to Inktomi

To submit a homepage using Add URL, you would have to use one of Inktomi's partners.

HotBot UK is recommended for this purpose.

HotBot UK Add URL - <http://www.hotbot.lycos.co.uk/submit.html>

The submission through the addurl feature is levied a penalty on ranking if that is the only way the page has been crawled. If the same page gets covered through normal crawling or through paid inclusion, the ranking penalty is removed.

Inktomi has a paid submission program called "Search Submit" that charges a fee of \$39 a year for your homepage. More web pages can be added for \$25 per web page. After a year, if you do not renew the service, Inktomi might drop your homepage. Inktomi's partners sell this program.

Search Submit - http://www.inktomi.com/products/web_search/submit.html

Inktomi allows maximum 20 web pages from each website to be submitted per day.

Inktomi also provides Bulk Program wherein you can add thousands of web pages at a time. A small fee is charged each time someone clicks on your web page listing.

4.5 Submitting to Fast Search

The Add URL page for Fast Search: http://www.alltheweb.com/add_url.php

With this it might take up to six weeks until your web pages show up.

Fast Search's paid submission service cannot be accessed directly by you. It can be used through its partner Lycos. The fee for this service is \$30 for one year. Additional pages can be submitted at \$12 per page.

Paid submission service – <http://search.lycos.com/searchservices/>

Fast Search also provides Bulk Program wherein you can add thousands of web pages at a time. A small fee is charged every time someone clicks on your web page listing.

4.6 Submitting to Teoma

This is owned by Ask Jeeves, one of the popular search engines. Teoma, however, doesn't have a free Add URL feature. Teoma, being a web crawler can always include your homepage for free if there are other web pages pointing to your homepage. Like other engines, it has a paid submission program for a fee of \$30 a year. Additional pages can be submitted at \$18 per page. Paid service can be accessed at:

Ask Jeeves/Teoma Site Submit - <http://ask.ineedhits.com/>

4.7 Submitting to AltaVista

This is another very popular search engine. AltaVista does have submission via the Add URL feature at <http://addurl.altavista.com/sites/addurl/newurl>. As discussed with other search engines, free submission might take more than a month until your web pages show up. An automatic submission through submission software may not work with this URL, as it requires insertion of a specific code that is displayed on entering the submission page.

Following services are offered by Alta Vista.

Express Inclusion

According to Alta Vista, this is the fastest method of including your URLs as Express Inclusion lets you add up to 500 URLs to AltaVista's search results within 2 business days. It does Daily refreshes (weekday only) and keeps your listings fresh.

Trusted Feed

This service is reserved for partners who submit more than 500 pages to the AltaVista database. Trusted Feed enables submission of custom titles, keywords and abstracts and provides the clients with highly qualified traffic through weekly updates.

Listing Enhancements

This service visually distinguish your pages in the AltaVista search results with Listing Enhancements. It also controls the look and feel of your results by adding logos, icons, custom taglines, and text links to your URLs.

Reseller Program

AltaVista's Reseller Program provides you with the opportunity to deliver highly qualified traffic to your clients' sites by using any of the following programs: Trusted Feed, Express Inclusion, and Listing Enhancements.

Basic Submit

This is the basic submit service which enables the clients to add or remove up to 5 URLs at a time from the AltaVista global database. URLs will generally be evaluated within four to six weeks of submission. Basic Submit is a free service. As discussed with other search engines, free submission might take more than a month until your web pages show up. An automatic submission through submission software may not work with this URL, as it requires insertion of a specific code that is displayed on entering the submission page.

AltaVista's paid submission program has a fee of \$39 per page for six months (\$78 for a year).

For additional 2 to 10 web pages the fee is \$58 per page per year, whereas for an additional 11 to 500 web pages the fee is \$38. AltaVista also provides Bulk Program wherein you can add thousands of web pages at a time. A small fee is charged each time someone clicks on your web page.

Fortunately, there are ways to verify and whether your site has been listed or not. This is described in the section on “verify and monitor listing”.

4.8 Submitting to Directories

Directories are very popular and are widely used by people as a source of information. Web crawler search engines may also have better chances of finding your website if it is listed with any of these directories. Most of the directories charge a fee for listing your website. Some sites such as Yahoo offer free submission as well.

An important aspect of submitting your website to a directory is to have a 25 word or less description of the website. This allows the web crawler search engines to efficiently find and include your website.

Yahoo, LookSmart and Open Directory Project are some of the most popular directory services on the World Wide Web.

4.8.1 Yahoo

Yahoo provides two submission options: Free submission known as “Standard” and paid submission known as “Yahoo Express”. The free submission cannot be used for

commercial websites. Also, with free submission there is no guarantee that your website will be approved for submission.

Submission of commercial websites can be done with Yahoo Express that has a fee of \$299 per year. Although a majority of websites are accepted in this category, note that, paid submission also doesn't guarantee acceptance of your website. It only ensures an answer whether your site was accepted or not. For non-commercial websites there is a one-time fee of \$299.

For submission of free non-commercial websites you will have to fill up a submission form that is displayed once you click the "Suggest a Site" link at the bottom of each category. In case of paid commercial websites you would have to fill up the submission form that can be accessed at <http://add.yahoo.com/fast/add?+Business>

4.8.2 LookSmart

LookSmart provides listings used by the very popular MSN Search service and AltaVista. Hence, like Yahoo, it is essential that you submit your website to LookSmart. LookSmart also provides two options for submission: Free submission for non-commercial websites and Paid submission for commercial as well as non-commercial websites.

LookSmart paid submission charges you a nominal fee every time a click is registered against your website. Free submission cannot be done directly from LookSmart. It has to

be done through a site owned by LookSmart known as Zeal.com. The sign-up form for this is available at

http://www.zeal.com/users/become_a_zealot.jhtml

Once your website is approved for submission, you can submit your website by using the “Manage” link at the top of each of the categories.

4.8.3 Open Directory

Having your website listed with Open Directory is essential. It provides results to Google, AOL, Lycos and Netscape Search. Open Directory doesn't have a paid submission service. Commercial as well as non-commercial websites can be submitted to Open Directory absolutely free of cost. This, however, has its drawbacks.

There is no guarantee if and when your website will be approved for submission. Having mentioned that, Open Directory does generate high traffic for your website, once approved. Thus, submission to this directory is worth the uncertainty involved.

Submission can be done by using the “Add URL” link at the top of each of the categories. Generally, if your website is accepted, it would appear within the directory in about three weeks' time. If rejected, there is no limit to how many times you can resubmit your website.

These are three most important Directories on the World Wide Web. Getting your site listed on each of these is a must as it can lead to a significantly large audience for your website.

4.8.4 Some tips for Submission to Directories

The maximum length of the Title of the website should be 60 characters for Yahoo, 65 characters for LookSmart and 100 characters for Open Directory.

Description of your website is essential for submission to any Directory. The maximum length of this description for Yahoo is 25 words or 200 characters, 170 characters for LookSmart and 25 to 30 words for Open Directory.

Yahoo requires that you submit the Title, Description, Your name and Email for website submission, whereas LookSmart and Open Directory only require the Title and Description of the website.

Yahoo allows a maximum submission for one category per website; LookSmart and Open Directory allow a maximum submission for one category per URL.

4.9 Submitting to the big ones

There are currently hundreds of search engines out there, most of them are not much more than an advanced FFA page. You can surely not submit your site to each of these. The best strategy to follow is to use automated submittal software as discussed earlier for the Search Engines that are less popular, while hand submitting to the top 10 Search Engines.

Beware of the ads such as 'submit to the top 500 search engines for only \$99', because generally only the top 10 will drive traffic to your site. It doesn't take much time to manually submit to these top engines and your \$99 could better be spent on overture.com or buying ads in ezines or whatever.

At the moment the top ten search sites - meaning both directories and search engines account for just over 93% of all search engine traffic. The other 6.something% is made up of hundreds of sites claiming to be search engines. Even at that the 11th - 15th biggest search engines make up most of that figure. So what are the search sites you need to concentrate on? Some of the top search engines are discussed above. There are a few others which are also quite popular. The following is a suggested list of search engines and directories.

- Yahoo.com

- Dmoz.com (ODP)
- Looksmart.com
- Google.com
- Alltheweb.com (Fast)
- Inktomi (AOL, Hotbot, MSN + more)
- Altavista.com
- Lycos.com
- Teoma
- Overture.com (GoTo.com) **Paid Inclusion**
- Directhit.com
- Askjeeves.com

5. Intelligent ways of improving page rankings

Experts feel that web pages have to be search engine friendly in order to improve search engine rankings. Having a well designed and relevant web copy is crucial. This will help search engines easily index your web pages and rank them high.

Besides internal page factors including frequency and positioning of relevant keyword phrases, the leading search engines now suggest the importance of off the page factors, particularly links from other domains.

Google ranks individual URLs based on which other URLs link to them, which URLs link to those, and so on. Thus the steps you have to take are:

Choose the terms that you wish to include as the most productive search terms for your site, terms that you associate with the correct target visitors to your site. Put those in the search engines and find what pages rank near the top for them. These top-ranking sites/pages would be on the top of your list, if you needed to plant links to your site on other pages.

Once you have found the top ranked pages/sites for your chosen search term(s), you should check what sites or pages link to that site. If you get links from any of these sites

to your target pages, you may be able to dramatically improve your ranking on Google and other crawler based Search Engines.

If a website has a lot of interlinked pages, then even an obscure page from that site could be a reasonably effective candidate for having a link to your site. It is not essential that the link should be on the home page or prominent page of a popular site.

Creating fake domains and letting those domains point to your site is a trick thought to be effective for improving page ranking at Google. However Google claims that this is not true and they are able to spot duplicate domains and domain scams.

5.1 Link Popularity

Your "link popularity" is simply a count of the number of web pages that are linked to you. Improving your website's link popularity is absolutely vital for improving the visibility of your website with regard to Search Engines. You may want to know your link popularity for two reasons. The first is that your link popularity will improve your ranking on all crawler search engines. As discussed earlier, all crawler-based search engines have a component called the spider, which crawls from one webpage to another through links. Hence, more the websites linking to your website, the better are your chances of getting listed through a search engine. The second reason is that you might want to know which websites are linked to you and potentially referring traffic.

Search engines give sites with good inbound and outbound links a higher ranking. The logic goes that if you provide outbound links to other material you are providing a valuable service: and, if other sites link to you then you must have content of value.

5.1.1 Link Analysis and measuring Link Popularity

The best way to discover how people are finding your web site is to analyze your site's activity logs. If you are unable to analyze their logs can instead use search engines to track down referral links. In particular, this method gives you an idea of how "popular" a search engine believes your site to be. Be aware that "popularity" is only one part of the link analysis systems that search engines such as Google use to rank web pages. The quality and context of links is also taken into account, rather than sheer numbers.

You can use link:'site URL' feature of many search engines to list all the pages that link to the selected site, and that too in order of Page Rank. For Google, North Light and AltaVista, use link:xyyzz.com to find the listing of pages that link to the web site www.xyyzz.com For Alltheweb use link.all instead of link and for Inktomi use 'linkdomain' instead of 'link' in the above example. The results would be a list of all pages (if indexed by the search engine) that link to your target site, listed in the order of popularity.

If you need to find the link to specific pages instead of to an entire site, then the above link: feature will not work. Use the Advanced search features offered by HotBot and

MSN Search, enter the full URL of the target page including http:// and use the option “links to URL” or similar.

Some sites offer to run comparison of the links to a chosen site vis-à-vis three other chosen sites. <http://linkpop.marketleap.com> and <http://www.linkpopularitycheck.com> are two such sites where you could submit your target URL and three other URLs that you wish to have a comparison done. www.linkpopularity.com is a site that will analyze the link popularity of a chosen URL in three prominent search engines.

Link analysis is somewhat different than measuring link popularity. While link popularity is generally used to measure the number of pages that link to a particular site, link analysis will go beyond this and analyze the popularity of the pages that link to your pages. In a way link analysis is a chain analysis system that accords weighting to every page that links to the target site, with weights determined by the popularity of those pages. Search engines use link analysis in their page-ranking algorithm. Search engines also try to determine the context of those links, in other words, how closely those links relate to the search string. For example if the search string was “toys”, and if there were links from other sites that either had the word toys within the link or in close proximity of the link, the ranking algorithm determines that this a higher priority link and ranks the page, that this is linked to, higher.

As a site owner, you want to seek links from good pages that are related to the terms you want to be found for. Linking strategy is not a trick as many get rich quick merchants

would have you believe. Links for the sake of links have no value whatsoever. Indeed, they can damage your rankings. So forget about link farms and other such nonsense. A small number of inbound links from great, relevant sites will be much more valuable than many links from low-traffic, irrelevant sites.

However, you should not become obsessed by link popularity alone. Treat linking as one important aspect of your Search Engine Optimization strategy. Decide how much time and effort you are prepared to invest in relation to your other activities and be disciplined about your approach. Monitor your results and adapt your strategy as necessary.

5.1.2 Where and how to seek reciprocal linking

Once you have found the candidate sites that have high link popularity and link quality as seen through link analysis, the next step is to choose those that you believe may agree to reciprocal linking. Your competitors obviously would not, but in respect of others, you must try. Several sites have a page where they list useful links or relevant links, as a service to their site visitors. You can locate the email of the company or of the person who handles link requests for such list pages.

When you make a request for reciprocal linking, approach the target site owner or webmaster or link request handler by providing your URL and a short description and explain how providing this link would be a valuable and useful addition for the benefit of their audience. The description is important as often that would be what appears in their

links page. Offer them reciprocal linking from your site's links and resources page. Better still tell them that you have already provided a reciprocal link to their site.

Build a good links and resources page on your website. Present the links in an organized manner so that this is useful to your site visitors too. Of course, you do not want this links page to be amongst the first pages accessed by your visitor, as this may induce the person to leave your site. Some webmasters try to build a standalone links page that is totally isolated from all other pages on their site. This would perhaps provide an URL to the reciprocal link provider, but in reality it is not a genuine link page at all. Avoid these tactics; they will not work in the long run.

Finally, not all sites are equal and therefore not all links are equal. A link from a high traffic industry portal is worth infinitely more than a link from a low-traffic free-for-all site. Concentrate on giving the search engines what they really want - great content, well-organized, well-published and linked to other relevant material. Concentrate on that and you will be rewarded.

5.1.3 Inward Linking

Like reciprocal linking, inward links to your website can be an effective strategy to increase your website's visibility to Search Engines. Inward links are links pointing to your websites from other websites without providing a reciprocal link from your website.

There are many techniques to improve inward linking. Many of these have enjoyed success with Search Engines. The most proven technique for inward linking is through Ebooks. You can offer interesting and educative Ebooks for free to other websites and they could install them on their sites. The Ebook you create would have a link to your website. This will allow a spider to crawl through that link and visit you website. For example, a footer on every alternate page can have a link to your website that would increase the probability of your website being listed with a crawler based Search Engine.

Other techniques include posting newsletters, white papers, news stories and press releases to other websites, particularly industry specific and general portals. The newsletters and press releases would contain a link pointing to your website, thus, increase its visibility to crawlers.

Affiliate programs also help in improving inward linking. In affiliate schemes, you provide incentives (usually a commission on the sale of your product or service) for other websites to become affiliates (i.e. *carriers*). These affiliates then generate you direct traffic; the added bonus comes in the form of those inward links to your site. Affiliate Programs create powerful alliances between your web site and your various "affiliate" web sites. Providing affiliate links to your website would improve your website's search engine ranking. It makes your website more visible to crawlers.

5.2 Meta tags and keywords

While Meta tags are not the complete answer to the question of "How do I improve my search engine ranking?" they can help with some search engines. Since there are millions of pages with Meta tags, you can add all the pages you want and still not control a sizeable percentage of the pages on the World Wide Web.

What are Meta tags? They are information inserted into the "head" area of your web pages. Other than the title tag, information in the head area of your web pages is not seen by those viewing your pages in browsers. Instead, Meta information in this area is used to communicate information that a human visitor may not be concerned with. Meta tags, for example, can tell a browser what "character set" to use or whether a web page has self-rated itself in terms of adult content.

Meta tags may help you with some search engines, so you'll want to consider adding them to every page you create. On the other hand, you can find many highly ranked web pages without Meta tags. For example, Meta tags have no effect on how humans will view your pages and enter your information into directories like Yahoo.

5.2.1 Tags that do matter

Titles

The TITLE tag is an important one for search engines like Google, as that is often the first one indexed and it is given higher weighting in the relevance rankings. You have to pay attention to this tag. Keep it short (lower than 40 characters) and let the tag have material relevant to the keywords used during search. As an example, if you had stuffed your Company Name as the TITLE tag, it would not help. The visitor that you want to attract is unlikely to be looking for your Company name in the search string. Think creatively as to what keywords people would use if they were looking for goods or services that your site offers.

Keywords and Description

Two Meta tags that are important are KEYWORDS and DESCRIPTION. You have to be very careful about how these are developed and positioned. Frequency and location of the keyword that is being searched are important criteria that determine relevance and hence page ranking. The search engine would generally consider the page more relevant if the keyword that is being sought is in the TITLE tag or is in the KEYWORDS tag near the top of the page. Similarly if the sought after keyword is found being repeated in the page, it may give the impression that this is a more relevant page and improve ranking.

There is a caveat, though. The above is only a general rule that has been often followed by many search engines, but then there are many variants to it. Several players in the SEO industry have tried to proclaim that this is the gospel truth and hence it has spawned a large number of experts that suggest and resort to keyword stuffing and spamming

(repeating long strings of keywords). The result can often be just the opposite. Some search engines penalize pages that have keyword spamming. Some will just ignore these pages. Some engines also do not read Meta tags. The intelligent method today is to stay away from spamming and to use tags judiciously. Blend your technique to attain the right frequency and location, but stay away from any excess or spamming.

You can provide an in-depth list of words and phrases in the KEYWORD Meta tag. These words should have some relevance to the specific page or, at least, to your website. While you can vary the case of the keywords, you'll want to concentrate on the lower case because over 90% of the searches either use lower case or are conducted on search engines that are not case sensitive.

These keywords should contain variations on the same theme. If your site was about gardening, you could use garden, gardening, home and garden, home gardening, vegetable garden, and herb garden. These are all words that might be used in searches for information that your site might provide. The keywords Meta tag is not intended to replace the actual text on your website. This tag is simply to aid the spider in collecting accurate information about your web pages.

The DESCRIPTION tag is used by search engines like Inktomi for the page summary that is displayed on the results page. This summary is what the visitor will read and decide whether he/she wishes to enter your site. If the description is just full of repeated

keywords, it won't do you any good, even if your page is ranked high. You still do not have a visitor. May be you have put off a visitor.

Meta Robots Tag

This is probably the only other prominent Meta tag. This is a very peculiar tag in the sense that it indicates what web pages should not be indexed by Search Engines. The Robots tag is inserted in between the header tags. An example of the Robots Tag is given below:

```
<HEAD>  
<TITLE>This page should be kept out of Search Engine listings</TITLE>  
<META NAME="ROBOTS" CONTENT="NOINDEX">  
</HEAD>
```

By default, a crawler will try to index all your web pages and will try to follow links from one page to another. This can be prevented by using the Robots Tag. Most major search engines support the Meta robots tag. The Meta robots tag also has some extensions offered by particular search engines to prevent indexing of multimedia content.

There are other Meta tags apart from the ones explored above but most of them are simply ignored by almost all search engines.

To sum up, some search engines will give you a boost if you have Meta tags. But don't expect that to necessarily be enough to put you in the top ten. Meta tags are mainly a design element you can tap into, a crutch for helping information-poor pages better be acknowledged by the search engines.

5.2.2 Choosing keywords that are focused and specific

This is an important step in the SEO exercise. What keywords do you need to emphasize and include? How do you choose the most relevant keywords that will be used by your target audience? Pose yourself the question: What would my target visitor be looking for and for which queries I would like to lure him/her to my site? Imagine those queries, as many of them as you can think. Apply this to all categories of visitors that you are targeting. Then list those queries and formulate your keywords from those.

Let your creativity develop grammar and synonym variants of these keywords. The same would apply to descriptions; however while keywords have to match closely with search strings, descriptions should be drafted to allure the visitor after he has seen your listing. It should tell the visitor that your site is indeed offering him best information or outcome for what he is looking.

Select phrases (at least two words), rather than single words, as there would be too many contenders for single word searches. You are much better off focusing on specific search keywords and using longer phrases.

5.2.3 Location of keywords

The location of keywords on your website is vital. You may have placed keywords at only a few locations but if they are inserted at important positions, your search engine ranking may be boosted. The most important position is the Title of the page. Ensure that your keyword is placed in the Title.

Another proven strategy is to place keywords towards the beginning of the web page. For instance, inserting keywords in page headlines, other sub-titles, and introductory paragraphs would certainly help rankings. Avoid using images or tables at the beginning of the page. This will only shove keywords down on the web page, and in turn harm your search engine rankings.

Amongst other strategies, “stemming” is another important technique for placing keywords. This is discussed in detail later in this chapter.

5.2.4 Making keywords relevant and consistent

Positioning Meta tags and keywords into web pages for better ranking should not be confused with spamming. Whatever you do, avoid spamming and keyword stuffing. Search Engines have mechanisms that can spot spamming and as a result ignore such web pages. The key is to insert just “enough” keywords into the web page so as to make the content relevant. Keywords have to be consistent with the entire content of the web page.

In other words, do not insert non – relevant keywords just for the sake of improving page rankings.

Many webmasters believe that adding more graphics to a web page would make it more attractive. This is a misconception. A web page should be rich with HTML text rather than graphics. This makes the site more relevant as Search Engines can easily scan text.

5.2.5 Stemming and other word forms

Search engines often look for variants of words from a stem and this is referred to as stemming. Thus ‘play’ can also lead to a search for ‘plays’ or ‘playing’ or ‘players’. Similarly singular and plural forms of words and case can lead to different results for some search engines. Some advisers suggest that Capitalization is a better bet in choosing keywords. However, the wiser counsel is that these variants do not make that great a difference and one should prefer lowercase. Most searches are made in lowercase. Moreover, many search engines are not case-sensitive.

5.2.6 Focus on important keywords: Finding top keywords

You may like to know what are the most popular search words and strings, so that you could choose top keywords related to your own site. Some resources that help in this are discussed here.

Place your subject term in Overture's free Search Term Suggestion Tool. The result will be a list of all search terms related to the word you inserted that were most popular at Overture. The list is in the order of popularity. Thus you know what most people are looking for. Wordtracker has a fee-based service that lets you do the same based on Meta search engines such as Dogpile and Meta Crawler. Supplement this effort through the Related Searches feature that is seen in many search engines. Sites such as AltaVista, Yahoo, HotBot and others have this feature. Place your subject or one of the target keywords in the search string on these. From the results page go to the section: "Others searched for" or "Related searches" that is available on many search engines and you will find the other related terms listed there.

5.3 Write a persuasive home page

In the field of search engine optimization (SEO), writing a strong homepage that will rank high in the engines and will read well with your site visitors can sometimes present a challenge, even to some seasoned SEO professionals. Once you have clearly identified your exact keywords and key phrases, the exact location on your homepage where you will place those carefully researched keywords will have a drastic impact in the end results of your homepage optimization.

One thing we keep most people say is that they don't want to change the looks or more especially the wording on their homepage. Understandably, some of them went to great

lengths and invested either a lot of time and/or money to make it the best it can be. Being the best it can be for your site visitors is one thing. But is it the best it can be for the search engines, in terms of how your site will rank?

If you need powerful rankings in the major search engines and at the same time you want to successfully convert your visitors and prospects into real buyers, it's important to effectively write your homepage the proper way the first time! You should always remember that a powerfully optimized homepage pleases both the search engines and your prospects. In randomly inserting keywords and key phrases into your old homepage, you might run the risk of getting good rankings, but at the same time it might jeopardize your marketing flow. That is a mistake nobody would ever want to do with their homepage.

Even today, there are still some people that will say you can edit your homepage for key phrases, without re-writing the whole page. There are important reasons why that strategy might not work.

Your homepage is the most important page on your web site

If you concentrate your most important keywords and key phrases in your homepage

many times, the search engines will surely notice and index it accordingly. But will it still read easily and will the sentences flow freely to your real human visitors? There are some good chances that it might not. As a primer, having just 40 or 50 words on your homepage will not deliver the message effectively. To be powerful and effective, a homepage needs at least 300 to 400 words for maximum search engine throughput and effectiveness.

One way to do that is to increase your word count with more value-added content. This often means rewriting your whole homepage all over again. The main reason to this is you will probably never have enough room to skillfully work your important keywords and key phrases into the body text of your homepage. This may not please your boss or marketing department, but a full re-write is often necessary and highly advisable to achieve high rankings in the engines, while at the same time having a homepage that will please your site visitors and convert a good proportion of them into real buyers.

Here is the acid test that will prove what we just said is right: Carefully examine the body text of your existing homepage. Then, attempt to insert three to five different keywords and key phrases three to four times each, somewhere within the actual body of your existing page. In doing that, chances are you will end up with a homepage that is next to impossible to understand and read.

One mistake some people do is to force their prospects to wade through endless key phrase lists or paragraphs, in an attempt to describe their features and benefits. The other reason they do that is in trying to please the search engines at the same time. Writing a powerful and effective homepage around carefully defined keywords and key phrases is a sure way you can drive targeted traffic to your web site and keep them there once you do.

If some people still say re-writing a homepage takes too much time and costs too much money, think of the cost of losing prospective clients and the real cost of lost sales and lost opportunities. In the end, writing a strong homepage that will achieve all your desired goals will largely justify your time invested and the efforts you will have placed in the re-writing of your homepage.

5.3.1 Setting a theme for your Homepage

We discussed the importance of the Homepage. This section presents a recommended layout for your homepage in order to make it as search engine friendly as possible. This is where you set the theme of your site. Let's suppose the primary focus of your site is about online education. You also have secondary content that is there as alternative content for those not interested online education. There is also other content that you would like to share with your visitors. For example, this might include book reviews, humor, and links.

The top of your homepage, as discussed earlier is the most important. This is where you set the keywords and theme for the most important part of your site, the thing you really want to be found for. So you might start off the top of your index page something like this:

(After your logo or header graphic)

1) A heading tag that includes a keyword(s) or keyword phrases. A heading tag is bigger and bolder text than normal body text, so a search engine places more importance on it because you emphasize it.

Heading sizes range from h1 - h6 with h1 being the largest text. If you learn to use just a little Cascading Style Sheet code you can control the size of your headings. You could set an h1 sized heading to be only slightly larger than your normal text if you choose, and the search engine will still see it as an important heading.

2) Next would be an introduction that describes your main theme. This would include several of your top keywords and keyword phrases. Repeat your top 1 or 2 keywords several times, include other keyword search terms too, but make it read in sentences that makes sense to your visitors.

A second paragraph could be added that got more specific using other words related to online education.

3) Next you could put smaller heading.

Then you'd list the links to your pages, and ideally have a brief description of each link using keywords and keyword phrases in the text. You also want to have several pages of quality content to link to. Repeat that procedure for all your links that relate to your theme.

4) Next you might include a closing, keyword laden paragraph. More is not necessarily better when it comes to keywords, at least after a certain point. Writing "online education" fifty times across your page would probably result in you being caught for trying to cheat. Ideally, somewhere from 3% - 20% of your page text would be keywords. The percentage changes often and is different at each search engine. The 3-20 rule is a general guideline, and you can go higher if it makes sense and isn't redundant.

5) Finally, you can list your secondary content of book reviews, humor, and links. Skip the descriptions if they aren't necessary, or they may water down your theme too much. If you must include descriptions for these non-theme related links, keep them short and sweet. You also might include all the other site sections as simply a link to another index that lists them all. You could call it Entertainment, Miscellaneous, or whatever. These

can be sub-indexes that can be optimized toward their own theme, which is the ideal way to go.

Now you've set the all important top of your page up with a strong theme. So far so good, but this isn't the only way you can create a strong theme so don't be compelled into following this exact formula. This was just an example to show you one way to set up a strong site theme. Use your imagination, you many come up with an even better way.

It's important to note that you shouldn't try to optimize your home page for more than one theme. They just end up weakening each other's strength when you do that. By using simple links to your alternative content, a link to your humor page can get folks where they want to go, and then you can write your humor page as a secondary index optimized toward a humor theme. In the end, each page should be optimized for search engines for the main topic of that page or site section.

Search engine optimization is made up of many simple techniques that work together to create a comprehensive overall strategy. This combination of techniques is greater as a whole than the sum of the parts. While you can skip any small technique that is a part of the overall strategy, it will subtract from the edge you'd gain by employing all the tactics

5.4 Using Clickthroughs and Affiliate Programs for better ranking

In affiliate programs, sites that send you traffic and visitors, have to be paid on the basis of per click or other parameters (such as number of pages visited on your site, duration spent, transactions etc). Most common contractual understanding revolves around payment per click or clickthroughs. Affiliates use tracking software that monitors such clicks using a redirection measurement system. The validity of affiliate programs in boosting your link analysis is doubtful. Nevertheless, it is felt that it does not actually do any harm. It does provide you visitors, and that is important. In the case of some search engines re-directs may even count in favor of your link analysis. Use affiliate programs, but this is not a major strategy for optimization.

5.5 Doorblocks

Several pages in e-commerce and other functional sites are generated dynamically and have “?” or “&” sign in their dynamic URLs. These signs separate the CGI variables. While Google will crawl these pages, many other engines will not. One inconvenient solution is to develop static equivalent of the dynamic pages and have them on your site. Another way to avoid such dynamic URLs is to rewrite these URLs using a syntax that is accepted by the crawler and also understood as equivalent to the dynamic URL by the

application server. The Amazon site shows dynamic URLs in such syntax. If you are using Apache web server, you can use Apache rewrite rules to enable this conversion.

5.6 Submit a “site map” page or a page with links to all inner pages

One good tip is that you should prepare a crawler page (or pages) and submit this to the search engines. This page should have no text or content except for links to all the important pages that you wished to be crawled. When the spider reaches this page it would crawl to all the links and would suck all the desired pages into its index. You can also break up the main crawler page into several smaller pages if the size becomes too large. The crawler shall not reject smaller pages, whereas larger pages may get bypassed if the crawler finds them too slow to be spidered.

You do not have to be concerned that the result may throw up this “site-map” page and would disappoint the visitor. This will not happen, as the “site-map” has no searchable content and will not get included in the results, rather all other pages would. We found the site wired.com had published hierarchical sets of crawler pages. The first crawler page lists all the category headlines, these links lead to a set of links with all story headlines, which in turn lead to the news stories.

5.6.1 Size of submitted page

We have written above that the spiders may bypass long and “difficult” pages. They would have their own time-out characteristics or other controls that help them come unstuck from such pages. So you do not want to have such a page become your “gateway” page. One tip is to keep the page size below 100 kb.

5.7 Submit only key pages

You do not have to submit all the pages of your site. As stated earlier, many sites have restrictions on the number of pages you submit. A key page or a page that has links to many inner pages is ideal, but you must submit some inner pages. This insures that even if the first page is missed, the crawler does get to access other pages and all the important pages through them. Submit your key 3 to 4 pages at least. Choose the ones that have the most relevant content and keywords to suit your target search string and verify that they link to other pages properly.

5.8 Working with Frames

Many websites make use of frames on their web pages. In some cases, more than two frames would be used on a single web page. The reason why most websites use frames is because each frame’s content has a different source. A master page known as “Frameset” controls the process of clubbing content from different sources into a single web page.

Such frames make it easier for webmasters to club multiple sources into a single web page. This, however, has a huge disadvantage when it comes to Search Engines.

Some of the older Search Engines do not have the capability to read content from frames. These only crawl through the frameset instead of all the web pages. Consequently web pages with multiple frames are ignored by the spider. There are certain tags known as “NOFRAMES” (Information ignored by frames capable browser) that can be inserted in the HTML of these web pages. Spiders are able to read information within the NOFRAMES tags. Thus, Search Engines only see the Frameset. Moreover, there cannot be any links to other web pages in the NOFRAMES blocks. That means the search engines won't crawl past the frameset, thus ignoring all the content rich web pages that are controlled by the frameset.

Hence, it is always advisable to have web pages without frames as these could easily make your website invisible to Search Engines.

Making frames visible to Search Engines

We discussed earlier the prominence of frames based websites. Many amateur web designers do not understand the drastic effects frames can have on search engine visibility. Such ignorance is augmented by the fact that some Search Engines such as AltaVista are actually frames capable. AltaVista spiders can crawl through frames and index all web pages of a website. However, this is only true for a few Search Engines.

The best solution as stated above is to avoid frames all together. If you still decide to use frames another remedy to this problem is using JavaScripts. JavaScripts can be added anywhere and are visible to Search Engines. These would enable spiders to crawl to other web pages, even if they do not recognize frames.

With a little trial and error, you can make your frame sites accessible to both types of search engines.

5.9 Using Robots.txt to your advantage

We discussed the ROBOTS tag in brief earlier. Let us understand this tag a little more in detail.

Sometimes we rank well on one engine for a particular keyphrase and assume that all search engines will like our pages, and hence we will rank well for that keyphrase on a number of engines. Unfortunately this is rarely the case. All the major search engines differ somewhat, so what's get you ranked high on one engine may actually help to lower your ranking on another engine.

It is for this reason that some people like to optimize pages for each particular search engine. Usually these pages would only be slightly different but this slight difference could make all the difference when it comes to ranking high.

However because search engine spiders crawl through sites indexing every page it can find, it might come across your search engine specific optimizes pages and because they are very similar, the spider may think you are spamming it and will do one of two things, ban your site altogether or severely punish you in the form of lower rankings.

The solution in this case is to stop specific Search Engine spiders from indexing some of your web pages. This is done using a **robots.txt** file which resides on your webspace.

A **Robots.txt** file is a vital part of any webmasters battle against getting banned or punished by the search engines if he or she designs different pages for different search engine's.

The **robots.txt** file is just a simple text file as the file extension suggests. It's created using a simple text editor like notepad or WordPad, complicated word processors such as Microsoft Word will only corrupt the file.

You can insert certain code in this text file to make it work. This is how it can be done.

User-Agent: (Spider Name)

Disallow: (File Name)

The User-Agent is the name of the search engines spider and Disallow is the name of the file that you don't want that spider to index.

You have to start a new batch of code for each engine, but if you want to list multiply disallow files you can one under another. For example –

User-Agent: Slurp (Inktomi's spider)

Disallow: xyz-gg.html

Disallow: xyz-al.html

Disallow: xxyyzz-gg.html

Disallow: xxyyzz-al.html

The above code disallows Inktomi to spider two pages optimized for Google (gg) and two pages optimized for AltaVista (al). If Inktomi were allowed to spider these pages as well as the pages specifically made for Inktomi, you may run the risk of being banned or penalized. Hence, it's always a good idea to use a robots.txt file.

The **robots.txt** file resides on your webspace, but where on your webspace? The root directory! If you upload your file to sub-directories it will not work. If you wanted to disallow all engines from indexing a file, you simply use the * character where the engines name would usually be. However beware that the * character won't work on the Disallow line.

Here are the names of a few of the big engines:

Excite - ArchitextSpider

AltaVista - Scooter

Lycos - Lycos_Spider_(T-Rex)

Google - Googlebot

Alltheweb - FAST-WebCrawler

Be sure to check over the file before uploading it, as you may have made a simple mistake, which could mean your pages are indexed by engines you don't want to index them, or even worse none of your pages might be indexed.

Another advantage of the Robots.txt file is that by examining it, you can get information on what spiders, or agents have accessed your web pages. This will give you a list of all the host names as well as agent names of the spiders. Moreover, information of very small search engines also gets recorded in the text file. Thus, you know what Search Engines are likely to list your website.

5.10 The Right Text Usage

5.10.1 Full Body Text

Most Search Engines scan and index all of the text in a web page. However, some Search Engines ignore certain text known as Stop Words, which is explained below. Apart from this, almost all Search Engines ignore spam.

5.10.2 Stop Words

Stop words are common words that are ignored by search engines at the time of searching a key phrase. This is done in order to save space on their server, and also to accelerate the search process.

When a search is conducted in a search engine, it will exclude the stop words from the search query, and will use the query by replacing all the stop words with a marker. A marker is a symbol that is substituted with the stop words. The intention is to save space. This way, the search engines are able to save more web pages in that extra space, as well as retain the relevancy of the search query.

Besides, omitting a few words also speeds up the search process. For instance, if a query consists of three words. The Search Engine would generally make three runs for each of the words and display the listings. However, if one of the words is such that omitting it does not make a difference to search results, it can be excluded from the query and consequently the search process becomes faster.

Some commonly excluded "stop words" are:

after

also

an

and

as

at

be

because

before

between

but

before

for

however

from

if

in

into

of

or

other

out

since

such

How You Can Finally Achieve Top Search Engine Rankings!

than
that
the
these
there
this
those
to
under
upon
when
where
whether
which
with
within
without

5.10.3 ALT Text / Comments

Search engines are unable to view graphics or distinguish text that might be contained within them. For this reason, most engines will read the content of the image ALT tags to

determine the purpose of a graphic. By taking the time to craft relevant, yet keyword rich ALT tags for the images on your web site, you increase the keyword density of your site.

Although many search engines read and index the text contained within ALT tags, it's important NOT to go overboard in using these tags as part of your SEO campaign. Most engines will not give this text any more weight than the text within the body of your site.

5.10.4 Invisible Text

Invisible text is content on a web site that is coded in a manner that makes it invisible to human visitors, but readable by search engine spiders. This is done in order to artificially inflate the keyword density of a web site without affecting the visual appearance of it. Hidden text is a recognized spam tactic and nearly all of the major search engines recognize and penalize sites that use this tactic.

5.10.5 Tiny Text

This is the technique of placing text on a page in a small font size. Pages that are predominantly heavy in tiny text may be dismissed as spam. Or, the tiny text may not be indexed. As a general guideline, try to avoid pages where the font size is predominantly smaller than normal. Make sure that you're not spamming the engine by using keyword

after keyword in a very small font size. Your tiny text may be a copyright notice at the very bottom of the page, or even your contact information. If so, that's fine

5.11 Search Engine Regionalization

Almost all Search Engines serve different countries. Search Engines do list content from other countries but most of the content that is listed is either US or UK dominated content.

With this in mind, most popular Search Engines have started deploying regional editions that serve only a specific country. For instance, Google has an Indian edition (<http://www.google.co.in>) that caters to the Indian audience.

Given below are some of the types of Search Engine Regional Editions.

5.11.1 Regional Interface

Regional Interface is nothing but a translated version of the main Search Engine. Many Search Engines have interfaces in different languages such as French, German, Spanish, Japanese etc. However, the only difference between these regional interfaces and the main version of the Search Engine is that the language used on the interface is not

English. In other words, if you search using a keyword on both the interfaces, the listings are exactly the same.

Regional Interfaces are aimed at an audience that does not understand English.

5.11.2 Human Categorization

Human Categorization, as the name suggests, is categorization of websites by human beings. Search Engine employees categorize different websites into regional listings. Websites that are more relevant to a specific country are listed in that edition of the Search Engine. Hence, for a French edition a search would mainly list documents from France. This eliminates the problem mentioned above. The only caveat being that the whole process is manual.

Directories such as Yahoo, LookSmart, and Open Directory make use of this process.

5.11.3 Domain Filtering

Domain Filtering automatically segregates websites from different countries into their respective regional editions. This segregation is done on the basis of domain names. For instance a website from Australia would generally have a domain .au. The Domain filtering mechanism looks at the domains of all websites and creates a country specific edition listing.

Some Search Engines also have region specific editions which contain listings from the whole of that region. As an example: A French edition of Google may also return German or Spanish websites in some cases.

Domain Filtering has a drawback though. This mechanism can only filter out websites based on the domain name, and hence .com is always considered to be a United States website. This is obviously not true. Many websites from other countries also have .com domains.

5.11.4 Domain Crawling

Domain crawling is probably the best solution for maintaining both a main site and a regional version. With domain crawling the regional listing is far more comprehensive as compared to the other mechanisms explained above. Some pages, although regional may be listed in the main listing as well.

5.12 Spamming – A Definite No-No

A couple of years ago spamming may have worked wonders for your website. However, with sophisticated algorithms being developed by all popular search engines, spamming

can only backfire. Algorithms, these days, can easily detect spam and not only ignore your website but also ban your website.

Besides, instead of spending considerable time and effort on spamming you can always follow other proven strategies and have a higher rank with most search engines. Spamming can also easily irritate readers. Think about it – if your homepage has unnecessary repetitions of a particular keyword, it is bound to frustrate a reader. Consequently your site, instead of being content rich, would be junk rich. This can have nothing but a negative impact on your business.

5.13 Search Engine Cloaking explained

Search engine cloaking is a technique used by webmasters to enable them to get an advantage over other websites. It works on the idea that one page is delivered to the various search engine spiders and robots, while the real page is delivered to real people. In other words, browsers such as Netscape and MSIE are served one page, and spiders visiting the same address are served a different page.

The page the spider will see is a bare bones HTML page optimized for the search engines. It won't look pretty but it will be configured exactly the way the search engines want it to be for it to be ranked high. These 'ghost pages' are never actually seen by any real person except for the webmasters that created it of course.

When real people visit a site using cloaking, the cloaking technology (which is usually based on Perl/CGI) will send them the real page, that look's good and is just a regular HTML page.

The cloaking technology is able to tell the difference between a human and spider because it knows the spiders IP address, no IP address in the same, so when an IP address visits a site which is using cloaking the script will compare the IP address with the IP addresses in its list of search engine IP's, if there's a match, the script knows that it's a search engine visiting and sends out the bare bones HTML page setup for nothing but high rankings.

There are two types of cloaking. The first is called User Agent Cloaking and the second is called IP Based Cloaking. IP based cloaking is the best method as IP addresses are very hard to fake, so your competition won't be able to pretend to be any of the search engines in order to steal your code.

User Agent Cloaking is similar to IP cloaking, in that the cloaking script compares the User Agent text string which is sent when a page is requested with it's list of search engine names (user agent = name) and then serves the appropriate page.

The problem with User Agent cloaking is that Agent names can be easily faked. Search Engines can easily formulate a new anti-spam method to beat cloaklers, all they need to do is fake their name and pretend they are a normal person using Internet explorer or

Netscape, the cloaking software will take Search Engine spiders to the non - optimized page and hence your search engine rankings will suffer.

To sum up, Search engine cloaking is not as effective as it used to be, this is because the search engines are becoming increasingly aware of the different cloaking techniques being used by webmasters and they are gradually introducing more sophisticated technology to combat them. It may be considered as unethical by Search Engines if not used properly.

5.14 Does separating Domain Name Keywords with hyphens affect rankings?

Domain Names play a huge role in search engine optimization for obtaining high search engine ranking. Type in any keyword into Google for example, and the chances are high that you will find that more than 80% of the first 10 sites contain that particular keyword in their domain names. Thus, domain name is one area that can be fully utilized (obviously taking into consideration that you are optimizing a brand new site from scratch and not an existing one).

It is also a good idea to include hyphens (-) between each of the keywords within your domain name. This tells a search engine spider that each word is a separate word, not one continuous word. The search engines which use keywords in the domain name as a part of their ranking formula will not be able to recognize keywords unless they are

separated by a hyphen (or a slash or underscore for sub-directories). Search Engines prefer the use of hyphens in domain names because they can produce more accurate search results by being able to recognize specified key words in your URL.

Keep in mind that a domain with words separated by hyphens will be harder for users to remember (and may also decrease the value of the domain). However, if keywords are used with hyphens the keywords may be interpreted as such by search engines thereby helping rank a site higher. If at all possible, try to register both forms of the name for example, websitemarketing.com and web-site-marketing.com.

5.15 Taking advantage of Alphanumeric Search Engine listings

To have a good web site marketing plan, you must realize alphabetical priority of domain names is still used by some search engines as a key factor in their ranking formula. Alphabetical hierarchy is even more important to web site marketing, because this method is used by directories, which strictly list sites in alphabetical order based on the results of the keyword(s) search.

Your web site marketing plan must acknowledge Yahoo!. Yahoo! is the number one directory and search site with 52.7 million different visitors each month, accounting for 69.1 percent of all Internet surfers. An astounding 53.4 percent of all search-related traffic comes from Yahoo!, or as much as half the traffic received to many sites. Many prefer to use Yahoo! because each site submitted to Yahoo! is human-reviewed, delivering more accurate search results for their visitors.

To include alphabetical hierarchy in your web site marketing strategy, realize that alphabetical priority does not only consist of letters, it includes numbers that can rank higher than an "A".

This is the order:

- 0 1 2 3 4 5 6 7 8 9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

This means that the domain "www.1-800-ABC-NEWS.com" will rank higher with the search sites that use alphabetical priority in their ranking formula, than just "www.ABC.com"

5.16 Click Popularity

We discussed about Link Popularity. Another factor influencing search engine placement in some search engines is Click Popularity. The number of users clicking on links to your page from the search results is counted. Pages that are frequently clicked will get a popularity boost.

Your site is awarded a certain number of points each time someone clicks your link from the search results. If your Web site already has a high ranking you will get fewer points compared to a low ranking site. This way all sites have an equal chance to get click through points awarded.

Don't be tempted to click your own link over and over again. Repeated clicks from the same IP will be detected. Clicking on your link and quickly returning to the search engine again might actually hurt your rank. The search engines will believe you did not find anything interesting at the page. That is not a good search engine optimization strategy.

How can you influence click popularity then? By putting some work into your page title and description Meta tag. These are the main factors influencing people's decision to click your link. High quality content will make visitors stay at your search engine optimized web site, and will stop them from quickly returning to the search engine.

5.17 Registering a keyword rich domain

Domain name registration is one of the most important things to consider when you are designing a site for high search engine placement. Here are some dos and don'ts of domain name registration.

Some Webmasters use shared domains or sub-domains available for free from popular Web hosting services, or some kind of free domain name redirect service. This might be cheap, but if you want a good search engine placement, it's not an option.

- Some search engines do ban free Web hosts because search engine spammers frequently use them for hosting duplicate sites (mirrors) and doorway pages.
- Sharing domains or IP addresses with spammers can get your search engine position penalized or your entire site banned. Note this statement from AltaVista:

"You could wind up being penalized or excluded simply because the underlying IP address for that service is the same for all the virtual domains it includes."

- Most search engines limit the number of submissions or number of listings for each domain. This will make it very hard to get your site indexed. Other sites on the same domain might already take all the available spots.
- If you do manage to get your site indexed, the search engine will have a hard time finding the "theme" for your site if you are sharing a domain with other sites on many different subjects. Pages are no longer ranked one by one, all content within the domain is considered.
- Without your own domain, you will be forced to start working from scratch again if the host goes out of business or if he decides to change your URL. Many Webmasters have lost their search engine positions, link popularity and Web traffic because of this.

Keywords in the domain name are crucial. It makes sense to put your primary keywords into your domain name.

- Separate multiple keywords like my-keywords-phrase.com instead of typing it all in one word: mykeywordphrase.com. This will make it possible for the search engines to understand your keyword phrases correctly.
- Keep in mind that Yahoo and some other search engines reject domain submissions with URL's in excess of 54 characters. You would be wise to stay under the **55 character limit** when choosing a domain name.

- Directories like Yahoo, LookSmart and ODP will not look for keywords in the text of your page, and editors will often edit keywords out of your title and description. This leaves your internet domain name as the single most important place to put keywords for your site.
- Too many dashes in a domain name might trigger the spam filters of some search engines.
- Yet another benefit of keyword rich domain names is in reciprocal linking. If the domain name keywords appear within the text of incoming links, you will get a major boost in ranking, especially in Google.

5.18 Web Hosting

Choosing a Web host provider for your site is no small decision. When you make a choice of a Web page host, you are actually placing your Internet business in the hands of your Web host provider. Choosing the wrong Web page hosting plan can cause permanent damage to your search engine placement.

Selecting a Web page host is like renting an office. You want to get it right the first time. It is always possible to move somewhere else, but it's no easy task to move your entire business - not to mention the damage that already could be caused by then.

We are going to look at some little known facts about Web page hosting and how it can effect your search engine placement.

Did you submit your site over and over again without ever getting it indexed in the search engines? Or was your site indexed, but is nowhere to be found when you type your search terms into the engines? Your Web host provider might be the cause of that. As AltaVista states:

"If being found via search engines is important to your business, be very careful about where you have your pages hosted. If the hosting service also hosts spammers and pornographers, you could wind up being penalized or excluded."

AltaVista is not the only search engine enforcing this policy against certain Web Hosting services. This is a common practice causing many sites to lose their rank. Choosing the right host for your Web hosting needs is of great importance if you really care about your search engine position.

Here is a short list of things to watch out for

Shared IP hosting

Many Web hosting services do not give out unique IP addresses to customers. The name-based system of virtual Web hosting allows multiple domains to be hosted by a single IP. This means several hundred domain names could all be using the same IP address. Are you sharing an IP address with people you don't even know?

If someone else, who is hosted on the same server as your site, gets banned on a search engine, then your Web site will be caught and banned as well. This happens on a regular basis causing serious problems for those who are affected.

Downtime

If your Web host provider cannot keep servers up and running 24/7, people will not be able to find your site. The search engine spiders won't find it either. They will assume your site is gone and remove it from the search engine index. You will have to resubmit and wait weeks or months for the spider to re-index your site. If your site already is indexed and ranking well, it might lose its position altogether.

No matter what web host provider you use, it is highly recommended that you sign up for a free site monitoring service like EasyMonitor. This guarantees that you're alerted within minutes when your site goes down. This enables you to get to work immediately to solve the problem and dramatically increase the time your site is up and running.

The speed and reliability of your Web host provider will depend on several factors.

Look for a web page host with

- A T3 connection (or better) close to a primary Internet backbone
- Safeguards against systems, network or power failure

- Several connections to the Internet (in case one goes down)
- Backup power supplies and
- "Uptime guarantee" with at least 95% uptime

No logs

Access to raw web log files is one feature often missing with low quality Web hosting services. Some Web host providers provide logs, but do not include referrer information. Others delete logs frequently, making it impossible to use them effectively. Make sure your Web hosting service gives you access to raw server logs, preferably in "Extended Common Log File Format", a standard that can be understood by most log analyzers.

The server logs will give you some valuable information for your search engine optimization work, including what search engines people are using to find your site, exactly which keywords do they type into the search box, and when the search engine spiders visit your site. This is crucial information. You cannot properly optimize your site without it.

Other features include daily back up of your data, high bandwidth and, of course, the quality of their support services. In addition, you must consider the features offered by the Web page host. Some features to consider are

- The space available for your site

- Monthly transfer limits
- POP mail server availability
- Support for CGI, SSI, Perl, databases and FrontPage 2000 extensions
- Secure server access (SSL)
- Anonymous FTP and
- Shopping Cart Software.

5.19 Caps or no caps?

Some Search Engines are case sensitive. This gives rise to a new dilemma for webmasters. Should they include keywords in both lower case as well as upper case, especially in Meta tags? Most webmasters actually do so. For instance, they would list the keywords as a set of all possible combinations of upper and lower case.

Webmasters may feel that this is the safest technique to avoid losing potential visitors. However, such repetition may be considered by the Search Engine to be Spam! After all, higher the number of words, higher the combinations.

The best way to get around this problem is to list keywords in lower case only. Most surfers always search in lower case. There are very few cases where surfers use capitals even if the word is a proper name. Of course, you may run a very small risk of losing a

few visitors who use upper case, but this risk is minimal. Besides, you do not want to be banned for Spam.

5.20 A few Search Engine myths

There is an abundance of search engine information available on the web- some of it valuable, much of it contradictory. Throughout the years some prevailing search engine myths have developed. Some of these myths are still encouraged. Others are based upon techniques that were effective years ago but no longer work. Still others come from simple misunderstandings that inevitably come with a relatively new medium. Here is a discussion on some of the most common myths around

Myth: Meta tags are the most important factor in search engine rankings.

Fact: Many search engines (most notably Google) ignore meta tags completely due to constant abuse by webmasters. The only importance placed on meta tags these days is actually the meta description tag, which will appear as the description for the corresponding page on engines that use inktomi data (such as MSN). Meta tags are virtually irrelevant in the ranking algorithms of the top engines- but many people continue to believe that they are the only optimization strategy that they need.

Myth: It's impossible to do search engine optimization in-house.

Fact: It often is done in house, and done effectively. This is typically when a large corporation hires in-house talent that is largely devoted exclusively to promoting the website. However, it is unrealistic to expect someone with many other job functions to do a credible job of SEO. Much of the skills are acquired through experience- and it isn't usually desirable to have someone "experimenting" with the company website (especially considering that certain techniques can get sites penalized on engines or banned outright).

SEO isn't rocket science, but it also isn't something that can be learned overnight. When deciding whether to outsource SEO or do it in house, it is important to consider the actual costs involved. Often, when the necessary hours it takes to pay someone to learn on the job are taken into account, it is cheaper to outsource (and the results are almost always better). Only a careful evaluation of your goals and resources can determine the best course of action for your company.

Myth: Sites must be constantly resubmitted to retain rankings.

Fact: This is a scare tactic popularized by various submission services and software companies. In fact, it is a waste of money to pay to have your site resubmitted once it is already listed in an engine's database. It will not hurt your rankings to constantly submit (or else people would submit their competitor's sites to get them penalized), but it will not help, either.

Myth: Search engine optimization is not as effective as "traditional" marketing.

Fact: In many ways, it is more effective. Companies often spend countless dollars on direct mail, television and radio advertising, and bulk email without a second thought. The common thread with each of these strategies is that the prospect is "approached" by the company, and that the company must reach a great number of people to find a few motivated prospects. On the other hand, search engines can deliver highly motivated prospects directly to your website- people who have already demonstrated, through their use of particular key phrases, an interest in your products or services.

5.21 Things to Avoid

Doorway or Gateway pages

Doorway pages are pages that have been designed for one purpose only, to rank high in the search engines for a particular keyword and then direct the arriving visitors further into the site. Usually their design is relatively simplistic and includes a lot of text, but contains very little graphics or other fancy effects.

Doorway pages focus on particular keywords and phrases targeted for search engines and serve as an entry page to visitors. These used to be a popular means to achieve high rankings with Search Engines, but recently, many Search Engines have started completely ignoring websites with doorway pages.

Because creating doorway pages used to be so easy, especially when using programs that can automatically generate hundreds or thousands of pages in a heartbeat by using a simple template and merely placing the keywords in the right places, many sites lost control and created enormous amounts of them.

The search engines initially tolerated this behavior, but as their databases started being filled with doorway pages they soon realized that these pages would quickly conquer the top ranking spots from normal content-rich pages. This would reduce the quality of their indices and lower the value of the search results, resulting in users abandoning the search engine.

Search Engines such as AltaVista began removing doorway pages from its index and in some cases even banning the sites that were using them. Other engines followed suit and soon it was safe to say that the age of doorway pages, at least in the form we used to know them, had passed.

Instead of focusing on doorway pages, you should focus on optimizing your actual content pages.

Keyword stuffing and spamming

Important keywords and descriptions should be used in your content in visible Meta tags and you should choose the words carefully and position them near the top and have

proper frequency for such words. However it is very important to adopt moderation in this. Keyword stuffing or spamming is a No-No today. Most search engine algorithms can spot this, bypass the spam and some may even penalize it.

Dynamic URLs

Several pages in e-commerce and other functional sites are generated dynamically and have ? or & sign in their dynamic URLs. These signs separate the CGI variables. While Google will crawl these pages, many other engines will not. One inconvenient solution is to develop static equivalent of the dynamic pages and have them on your site. Another way to avoid such dynamic URLs is to rewrite these URLs using a syntax that is accepted by the crawler and also understood as equivalent to the dynamic URL by the application server. The Amazon site shows dynamic URLs in such syntax. If you are using Apache webserver, you can use Apache rewrite rules to enable this conversion.

Re-direct pages

Sometimes pages have a Meta refresh tag that redirects any visitor automatically to another page. Some search engines refuse to index a page that has a high refresh rate. The meta refresh tag however does not affect Google.

Image maps without alt text

Avoid image maps without text or with links. Image maps should have alt text (as also required under the American Disabilities Act, for public websites) and the home page should not have images as links. Instead HTML links should be used. This is because search engines would not read image links and the linked pages may not get crawled.

Frames

There are some engines whose spiders won't work with frames on your site. A web page that is built using frames is actually a combination of content from separate "pages" that have been blended into a single page through a 'frameset' instruction page. The frameset page does not have any content or links that would have promoted spidering. The frameset page could block the spider's movement. The workaround is by placing a summary of the page content and relevant description in the frameset page and also by placing a link to the home page on it.

Tables

When you use tables on the key pages and if some columns have descriptions while others have numbers, it is possible that this may push your keywords down the page. Search engines break up the table and read them for the content the columns have. The first column is read first, then the next and so on. Thus if the first column had numbers, and the next one had useful descriptions, the positioning of these descriptions will suffer. The strategy is to avoid using such tables near the top of the key pages. Large sections of

Javascripts also will have the same effect on the search engines. The HTML part will be pushed down. Thus again, place your long Javascripts lower down on key pages.

Link spamming

Realizing the importance of links and link analysis in search engine results, several link farms and Free for All sites have appeared that offer to provide links to your site. This is also referred to as link spamming. Most search engines are smarter to this obvious tactic and know how to spot this. Such FFA sites, as they are known, do not provide link quality or link context, two factors that are important in link analysis. Thus the correct strategy is to avoid link spamming and not get carried away by what seems to be too simple a solution.

6. Paid Placement Programs

Many Search Engines have now started offering paid placement programs in addition to their already existing programs. When someone searches for a specific term, the listings show up in the order of highest bidder first, second highest bidder second and so on. If two companies bid the same amount, the first bid gets priority for the listing and will be listed first. Such programs are called Paid Placement Programs or Pay-per-click programs. You only pay when someone clicks on your listing and is delivered to your website.

Benefits of Paid-Placement / Pay-per-click Campaigns

- Search Engine Marketing campaigns are effective for increasing web site traffic quickly.
- Pay-per-click programs will allow you to quickly test the ability for your web site to convert visitors to a call to action or make a purchase.
- You will be able to identify which keyword phrases will provide the best conversion rate.

While ordinary search results frequently include links to non-working pages, home pages with irrelevant content and filler listings, paid search content ensures a user will gain quick access to relevant sites of interest. Since each listing is displayed by either a

keyword search or relevant text link, you are ensured only visitors with interest are being sent to your site. Thus, pay-per-click programs offer a more targeted approach.

The most prominent and successful Paid Placement programs are Google Adwords and Overture.

In 1998, Overture was the first to charge advertisers to be listed in its search results. Now, all the major search companies have jumped in. So-called paid-search revenue grew 40% last year, to \$1.4 billion. It's now up to 23% of the \$6 billion Net advertising market, which shrank 17% last year, according to analysts' projections.

Placing ads near search results offers the simple appeal of the Yellow Pages, but with different economics. Overture made \$78 million last year on \$668 million in sales, about half of the paid-search market. That's more than Yahoo made on \$953 million of revenue, largely from traditional online ads such as banners. Paid search also pushed search specialists LookSmart Ltd. and Ask Jeeves Inc. into profitability in the fourth quarter. And the biggest potential initial public offering of an Internet company this year is Google Inc., which began offering pay-per-click search last year. Search ads will make up the lion's share of Google's revenue, estimated at \$350 million to \$400 million.

6.1 Google Adwords

Google offers a paid placement program called Google Adwords. Paid listings in Google appear above and to the side of its regular results. Like most Paid Placement or Pay-per-click programs you are charged a fee only when a visitor clicks on your listing. There is a \$5 activation fee for this service and there is no monthly minimum fee.

Some advantages of Google Adwords are:

- Your price is automatically lowered to one cent more than your closest competitor.
- No one can lock in the top position because user click through rates and CPC determine where your ad is shown. The most relevant ads rise to the top.
- Google offers a unique set of tools to forecast your budget and select target keywords.
- You can target your ads to users in a specific country or only to speakers of a specific language.

Features of Google Adwords

- **Competitive Pricing:** AdWords Discounter automatically monitors your competition and lowers your CPC to one cent above theirs.
- **Ranking:** Rank is determined by combination of CPC and click through rate. If an ad is irrelevant to users, they won't click on it and it will move down the page. Your relevant ads will rise, at no extra cost to you.

- **Campaign Management:** Multiple ads per keyword. Google tracks individual ad performance to help you monitor which ones perform well. AdWords Discounter takes care of your campaign for you, dynamically raising and lowering your CPC within the range you specify to keep you in the position you want.
- **Timing:** Ads go live almost immediately once you create them.
- **Minimum Deposit to start:** \$5.00 total startup cost
- **Monthly Minimum Fee:** None
- **When Billed:** Billed only for clicks received, after they're received.
- **Budgeted Delivery:** Set daily budget for your campaign. Continuous delivery each day over the time period at that level.
- **Target by Country or Language:** You choose who should see your ads from among 250+ countries and 14 languages.
- **Tools:** Enhanced keyword tool suggests other phrases AND most popular synonyms based on billions of searches. Results in better targeting and higher click-throughs.
- **Ad Groups:** Set a maximum price for an individual keyword or a group of keywords.

6.2 Google AdSense

This is a very recent tool from Google that has been earning rave reviews from experts all over. This tool focuses on a relatively new concept of marketing known as Contextual Marketing.

Contextual Marketing

This is a newly emerging but extremely targeted and effective form of promoting your online business. Consumers download specific software from a contextual marketing network organizer at no cost and install it on their desktops. When the consumer is about to make a purchase for a specific product or service or is placing a search query for that item, the software will beam advertising messages from suppliers of that product or service, offering some incentive to the consumer such as discounts or promotion offers. Through this facility, the advertiser (supplier) is getting an opportunity to promote his offering to a targeted consumer and that too at the point of sale or when there is an intention to buy. Thus contextual marketing offers a highly targeted marketing tool that connects the supplier with a genuine lead and spurs transaction.

The effectiveness of contextual marketing can be seen from the fact that response rates and conversion rates are from 5 to 20 times higher than the results from conventional online advertising.

Here's how Google AdSense applies this concept.

Google AdSense is a fast and easy way for website publishers of all sizes to display relevant, text-based, un-obtrusive Google AdWords ads on their website's content pages and earn money. Because the ads are related to what your users are looking for on your

site, you'll finally have a way to both monetize and enhance your content pages. The program is free, and Google pays you for clicks on the AdWords ads on your site.

There are no strict criteria for acceptance into the AdSense program, which, unlike other ad networks, does not place minimum traffic requirements on applicants. The only real criteria are the standard “acceptable content” requirements that exist just about everywhere. Of course, AdSense wants to attract quality content sites, and will only allow AdSense members to serve one ad per page – the service can’t be used for both banners and skyscrapers.

Once you’ve been accepted, you’ll be able to run AdSense advertisements on any site you own using the same ad code, provided you obey the guidelines. Reporting doesn’t occur in real time, but is updated regularly throughout the day. Currently, you cannot view reports based on a domain or site basis if you run the ads on more than one site.

This is what Google says about what makes AdSense different from others

This innovative new Google business solution will help you unleash the true revenue potential of your site.

AdSense Advantages

- AdSense is simple to join without any entry costs
- No technical skills required. Just a few lines of code to be copy pasted

- Only ads relevant to your content appear
- AdSense automatically shows appropriate ads for different sections. So it saves great deal of time and energy searching for relevant advertisers for different sections
- You're free to concentrate on providing good content and Google does the work of finding the best ads for your pages
- You can filter up to 200 URLs, so you can block ads for sites that don't meet your standards. You can also block strong competitors
- Access base of 100,000 advertisers

AdSense disadvantages

- Low flexibility -- Unlike other advertising sites where you can choose your advertisers and make choices on basis of payout, display and frequency, AdSense proves to be pretty inflexible.
- Free advertising -- The ad panels say "Ads by Google" - free advertising for Google. You don't earn anything if someone clicks on that link.
- You can filter out 200 URLs. For big sites this figure is not adequate
- Minimum payout is \$100, which is fair enough for most webmasters but keeps the smaller ones out.

· Revenue sharing isn't disclosed by AdSense. Considering Google's reputation maybe it can be trusted, but such disclosures especially with regards to revenue sharing should have been made.

Very recently, the company added new features to Google AdSense, including preset color palettes to coordinate colors for the Web sites and the ads, automatic user feedback mechanisms, and reporting tools that tell participants their click-through rates and how much money they've made at any given time.

Current AdSense participants include the Food Network, Weather.com, ABC.com, Internet Broadcasting Systems, Lycos Europe, New York Post Online Edition, Reed Business Information, US News & World Report, and iVillage.

6.3 Overture

When you advertise in Overture Premium Listings™, your business appears in the top U.S. search sites: MSN, Yahoo!, InfoSpace, Lycos, AltaVista, and Netscape.

Overture offers two sign-up plans, Fast Track or Self Serve. You have to select a few keywords with reference to your website. You have to then bid an amount (per click) and position in the search results for each of the keyword. If you have the highest bid for that position, Overture will display your website in the search result of the keyword selected, at the position selected, in each of the above mentioned search engines. Note that even if

your website is displayed in the search results you do not get charged the amount bid by you per click. It is only when a user clicks on your website link from the search results that you get charged that amount.

The higher the position, the higher is the current bid amount. Also, in order to keep your account active, you have to spend at least \$20 a month. Your account with Overture will be discontinued if the total clicks on your website amount to less than \$20.

The two sign-up plans, Fast Track or Self Serve, give you the choice of expert assistance or starting out on your own. Fast Track charges an additional one-time service fee of \$199 whereas Self Serve does not have any service fee.

When you sign up with Overture using the Fast Track Sign-up program, you will receive a customized proposal prepared by one of the Overture experts. When you receive your proposal, all you have to do is approve Overture's choices to get your search listings online. You can also make changes in the proposal yourself.

Two basic concepts apply to using these types of search engines (Overture and Google AdWords) to your best advantage:

- While you obviously want to pay the least possible amount for each hit you get, you're going to have to bid more if you want a higher listing and, therefore, more traffic. You have to carefully track the traffic you get from the search engine so

you know the value of that traffic. If you don't know the value, you could be paying too much for your traffic or missing great opportunities to generate more traffic if you can afford the higher price.

In other words, let's assume you're selling an item with a \$25 profit margin. If one (1%) percent of visitors to your site buy the product, then each visitor is worth twenty five (\$0.25) cents to you. If you get the traffic for less than that, you're making money on each sale.

- The other concept involves bidding for several hundred keywords at low prices such as one penny each. This way you get listed all over the search engine. Each keyword won't bring you much traffic by itself, but the total may be a hundred or more hits per day. If you're only paying a couple of pennies per hit for this traffic, then it should be profitable for you.

If you search popular terms like "business", you'll find the cost to be number one is quite high and the position is usually held by a large enterprise like Inc. magazine or Entrepreneur magazine. Large companies are often willing to invest millions of dollars in building their brands so they'll pay more for traffic than it would normally be worth. I strongly suggest you don't compete with them unless you are absolutely sure your business model will justify it.

Often, you'll be able to find a top ten ranking for a third or less than the price to be number "One". Since you're limited to buying only traffic that's profitable for you, this position will be better suited for you. Even if you end up far lower in the rankings, you must stay within your proper range for investment. You'll probably get less traffic than the number one ranking, but you'll get traffic you can make money on. That's far more important.

On another note, Overture is also planning to come up with its own Contextual Advertising business solution soon to compete with Google AdSense.

6.4 Searchfeed

Searchfeed is another pay-per-click program that has gained popularity in recent times. Based on site relevancy, you bid for placement within search results distributed to leading Web portals, industry specific sites and regional Internet service providers through the Private Label Feed (PLF) program. While the program is much like any other pay per click program, the difference is that Searchfeed has tied up with several portals and high traffic sites and visitors to those sites may thus use the Searchfeed engine to find suppliers.

You select relevant keywords for your website, create a title and description for each listing and bid on select keywords based on site relevance. You pay for each targeted lead only.

You determine the value of each unique visitor to your site using the most appropriate search terms to build continuous traffic. The more targeted the pay per click search term, the higher the bid you will want to make to ensure high placement. Because you can place an individual bid on each specific keyword or phrase being displayed, the value of each lead can be determined based on how targeted the interest of a potential visitor might be.

A deposit of \$25 is required to activate your account. Bid amounts can be as low as 1 cent.

6.5 Ah-Ha

Ah-Ha.com, Inc. applies unique advertising programs to join websites to potential customers. Founded in 1999 as a filtered search engine, ah-ha now offers an array of products and services to advertisers and webmasters throughout the Internet, including Distribution and Partner Programs, Paid Listings, LogoLinks™, TopLinks Creatives, Guaranteed Inclusion, and Skyscraper advertisements. The company's programs are designed to deliver targeted traffic to websites. The advertiser only pays for traffic delivered to their website and the partner builds bottom line revenue. The following programs are offered by ah-ha network.

Paid Listings - Paid Listings is an effective way to drive targeted visitors to your website. With this program, your company's title, description, and URL appear whenever

users type in keywords you selected, and the best benefit is that you only pay for visitors who click on your link.

LogoLink Account - With a LogoLink account, you have the advantage of our proven Paid Listings product, and you improve your clickthrough chances with your company logo displayed next to your listing. You have the bonus benefit of a graphic, which generates 4 times the clickthrough, and the branding benefits are endless.

Skyscraper Ads - Skyscraper ads are one of the most popular types of advertising used on the Internet today. ah-ha.com offers keyword-targeted, categorical, CPC, and run of network skyscraper ads that can benefit your company. Your banner will be prominently featured through our distribution partners including AltaVista, EarthLink, and LookSmart.

Guaranteed Inclusion - Guaranteed Inclusion allows you to submit your website listings for inclusion in the search results throughout websites such as Excite, DogPile, Verizon, WebCrawler, and more. This program gives you a yearly subscription to websites with an 85% Internet reach giving you an opportunity to reach millions

6.6 Sprinks

Sprinks is a self-service tool that provides advertisers an opportunity to place targeted ads across the web on a high-quality network of web properties. With Sprinks, advertisers can reach customers across a variety of online destinations.

Advertisers bid for placement within relevant search results (KeywordSprinks™), content pages (ContentSprinks™), opt-in, topical newsletters (DirectSprinks™), and geographically targeted content pages (LocalSprinks™). Sprinks' listing engine distributes those listings across a network of websites, putting those ads in front of the most appropriate audience available.

Advertisers are ONLY charged their bid amount when an interested customer clicks on a Sprinks listing and visits the advertiser's web site. You can open a Sprinks account for as little as \$50. Every dollar will go towards your campaign. There is a minimum bid of \$0.05 for KeywordSprinks™, and \$0.20 for ContentSprinks™, LocalSprinks™ and DirectSprinks™ listings. On some occasions, Sprinks will discount minimum bids for particular topics or assign a premium minimum bid to special topics.

6.7 Gaining more value out of pay-per-click programs

6.7.1 Determine the value of a visitor to your website

Determining the value of each keyword, and in turn each visitor, is vital. The simple formula goes like this: Divide the average number of new customers each month (a long

enough period to be statistically significant) by the average number of monthly visitors to get the percentage of visitors who actually become customers. If you multiply this percentage times your average profit margin on sales to new customers, you get a good idea of how much a visitor is worth to you on the first visit.

For example: If I average 10,000 visitors per month to my website and sell 150 of them a product with a \$20 profit margin, then I can easily calculate the percentage ($150/10,000$) to be 1.5%. Multiplying the 1.5% by \$20 yields a \$0.30 value for each visitor.

It's interesting to look at the bigger picture of this example as well. While making \$0.30 per visitor doesn't seem like much, the dollars get big quickly when you multiply pennies by thousands of visitors. My profits for the month with 150 sales would be \$3,000 if I had no advertising costs. If I can generate the visitors for a cost of \$0.10 each for a total cost of \$1,000, then I have a profit of \$2,000.

6.7.2 Submit targeted keywords

Submitting an extended list of keywords may be not be useful. In fact, you won't get as many potential customers to your website with an extended list of keywords as you would with focused and targeted keywords. Besides, you would also end up paying more without any benefits.

Targeted keywords would ensure that only those visitors who are actually interested in your product or service click on the search listing and visit your website. Thus, there are more chances of a visitor becoming a loyal customer.

6.7.3 Develop a catchy title and description

When the search engine user gets the result of a search, they'll see the highlighted titles for each listing followed by a short description. You should think of this just like a classified ad with a headline and short copy.

Here are some tips for formulating effective titles and descriptions:

- Use an attention getting headline to get as many people as possible to read the description. Here's where you want to be sure to use the important proven successful terms like "free" and "you".
- Write a description that succinctly tells the reader why he or she would be a fool not to click on your link. For most keywords this is the best plan because you're only paying a penny or two for the click-through. But, you'll want to be more selective about the copy for the more expensive bids. If you're going to pay a quarter for each click through, then you may want to have higher quality traffic. You don't want just anyone to click through. You want people more likely to buy.

- You should have the ability to track each keyword so you can determine the success ratios of clickthroughs to sales. All programs discussed above will allow you to use their system to track clickthroughs.

6.7.4 Re-evaluate and adjust your bids

If you determine that the traffic generated by a specific keyword is more valuable to you than other keywords, you can go online and immediately raise your bid for that term. You will instantly increase your traffic.

It would be a good idea to even track the patterns of which days and times generate the most sales from Search Engine traffic. During this prime time, you should raise your bids to increase traffic. As soon as the slot ends, you should lower your bids to the old levels.

6.7.5 Bid on as many low-cost relevant terms

The cost of a top position with any of the big PPC's depends upon the keyword you are bidding on and how many people search for it. If you have to bid on a popular word such as 'Internet marketing' be prepared to pay around \$2 - \$4 for a decent position. But what if you went for "Internet marketing articles" it costs less than a quarter of the price, it will produce far less visits of course, but is much more cost effective. Imagine doing this on a wide scale basis, securing lots of low-cost positions, the combined traffic from these

positions will add up to the level produced by a popular keyword - but at a fraction of the cost.

6.8 Disadvantages of Paid Programs

Disadvantages

Costs

The primary disadvantage of paid inclusion is the cost, although this factor naturally depends on your means. Typically, submitting to a few paid programs would cost anywhere around \$1000 a year.

Reach

A second disadvantage, perhaps more accurately described as a limitation, is that Google does not offer paid inclusion (and maintains that it never will). Since Google currently provides the primary results for three of the top four engines (Google, Yahoo, and AOL), engines that offer paid inclusion may only account for a fraction of your overall site traffic. Google typically updates about once every month, and there is no way to add your pages to their index any faster- which means that you will be waiting for Google to index your new (or newly optimized) pages regardless of which paid inclusion programs you use. Only after Google lists your pages will they appear in Yahoo and AOL results.

Considerations

There are many factors to consider when examining paid URL inclusion. The following five are some of the most common:

Are my pages already in the index?

Just because you can't find your pages when you enter search terms does not mean that your pages haven't been indexed. To see if your pages have been indexed, go to the engine and search for each of your exact page URLs in quotes. If each page shows up for the URL search but not for a search of any key phrases related to the page, paid inclusion will not help your rankings (your pages are already in the index and have been ranked according to their perceived value). It would be much more beneficial to invest some time and/or money in optimizing your pages for better rankings (you can still consider paid inclusion afterwards if you don't want to wait for the spider to revisit).

Is it a good investment for me?

Naturally, budgetary constraints can be a primary consideration. If you can't afford paid inclusion, then it obviously isn't an option. However, simply because you can afford it does not mean it is a good investment. For example, a business that sells a very inexpensive product online that is counting on volumes of traffic may not see a good return on their investment (again, 3 of the top 4 engines do not offer paid URL inclusion).

On the other hand, if your business has a high average dollar sale and you put a high value on each quality lead, you might consider immediate paid URL inclusion a no-brainer.

Do my pages change frequently?

If your web pages are subject to daily or weekly changes in content, paid inclusion may offer some additional benefits. When your pages are spidered frequently, all new content is indexed by the engine soon after it is added to your pages. This means that your pages will begin to appear in searches for terms related to the new content much more quickly.

Are my important pages dynamically generated?

Some search engine spiders have a problem finding and indexing pages that are dynamically generated (such pages often have a question mark somewhere in the URL). By paying to include the important pages of your dynamically generated website, you can be sure that they are in the engine's index, even if the "organic" spider would never find them on its own.

7. Verify and monitor listing and ranking

7.1 Spider spotting

The effectiveness of your efforts in submitting your pages for listing on search engines can be monitored and evaluated by two methods: spider spotting and URL check.

Spiders from search engines that visit your site and crawl pages leave some unique trace marks in your access log. This can tell you whether a spider has visited or not, what pages they have visited and also the frequency or duration of their visit.

The best way to identify spider visits is by finding out which visitors asked for the file robots.txt from your site. Only spiders make such a request, as this file is an indication to them to avoid covering the page in question. So the first thing a crawler would do is to check for this file. If you see the access log and analyze it using some convenient software, you would be able to spot all the visits that were initiated with this request. Then one can spot the host name and relate that to major search engines. Host names are related to the search engine company's name (it is the name of the site that hosts the spider). Another name that is used to identify such visits is the agent or browser names used by respective search engines. Get a list of host names and agent names from available resources (these names tend to change often) and also develop your own intuitive list by searching your access logs for all occurrences of known engine, host or

agent names. Concentrate only on the top engines; though you may find several other smaller and less known search engines visiting your site.

Pay attention to not only the total number of visits but to the activity pattern for each of the recent visits to actually judge how many pages they covered. This is a very good way of ensuring if submissions have worked or if other inducements such as links from other sites have worked or not. This also helps you to distinctly evaluate the effectiveness of submission, indexing and page ranking characteristics of your site.

Some examples of hostnames and agent names are as below:

- AltaVista: hostname may have altavista.com within its name; agent is often called Scooter
- Excite host name may have atex or excite.com and agent name is Architextspider.
- Inktomi agent and host names have inktomi.com and Slurp is often used as the agent name.
- Lycos uses lycos.com within its host name and Lycos Spider is often part of the agent name.

7.2 URL Search

One can use specific search strings in most search engines to find if your URL is included in their index and also to see how many pages are thus indexed. These search strings have been identified and compiled by some useful resources on SEO.

For searching the pages from your URL in **Google** for example, insert the following search string in Google search:

allinurl:yourcompanyname.com/webmasters/meta.html (this depends on the index pages of your site). In the Yahoo directory use the command u:yourcompanyname.com to find the listings for this URL. There are similar, but specific search strings applicable to each search engine.

Again checking the search engine for your URL is a good way to check what that search engine has indexed. Thus through spider spotting and URL checking you have evaluated and confirmed the effectiveness of your actions on submission and indexing activities.

8. Measuring Performance – Web logs analysis

The final step in your search engine optimization campaign is web log file analysis. The value of raw web logs are often not understood, but the truth is that your web log files are one of the most valuable resources you have as an Internet marketer.

Here are just a few examples of what your server logs can tell you:

- Measure the effectiveness of banner ads and PPC search engines.
- What search engine spiders have found your site?
- What keywords are people using to find your site?
- What search engines brings you most traffic?
- How many have book-marked your site?
- How well do people like your site?
- How many people visit your site?
- What pages are most popular?
- Why do people leave your site?
- Who is linking to your site?

Web log files tell you if search engine optimization and other marketing campaigns are working. They will show you exactly in what areas you have success and where you need

to put in more work. So, if you run a site but haven't seen your servers web log files, get a hold of them right away. They will teach you things you'd never expect about how to make your site better, not just for you, but also for the people who matter - your customers.

Understanding your customers

Logs are the best tool available for understanding your customers. Know how your customers find your Web site and why they are looking for it. This info will give you an idea about the mindset of your arriving visitors.

Log files can tell you not only what sites your visitors originate from (whether your visitor arrives via a Search Engine or from a "regular" site), but also what keywords they enter (in the Search Engines) to find you. The keywords give you an idea of your prospects' level of interest, as well as the kinds of benefits they seek.

Are you successfully drawing visitors through your site? A quick look at your log files will tell you. Of the people arriving to your site, how many are pulled through your page joint to the next page (or stage)? Once you know how visitors act on your site, you will understand how to change your Web site in order to improve your Conversion Rate (CR).

How You Can Finally Achieve Top Search Engine Rankings!

9. Review of SEO software and submission services

The use of automatic Submit tools for the major search engines and directories may not be desirable. If you submit to some search engines incorrectly your pages could actually be deleted permanently from their index or directory. At the same time, you may find manual submissions too time consuming. Experts have suggested that you use software that submits pages as if you were doing it by hand. TopDog, Web Position Gold and Submit Wolf Pro are good choices. Some others are: CommandoPro Submission Software, Add Web Site Promoter, VSE Be Found (Mac), Submission 2000.

Be very careful in selecting the company or service that you hire for search engine optimization, if you do not wish to do it yourself. Beware of those who promise top engine rankings through the use of Meta tags or by building doorway pages or using link farms or cloaked pages (submitting pages that are different from the actual live pages). If the firm tells you that you need to improve content, position keywords carefully, and seek quality reciprocal links, perhaps that is an indication of a good firm to work with.

10. ROI in search engine optimization

We have looked at various strategies and options related to search engine optimization. The key determinant that you have to use to formulate your own strategies would be to ascertain the ROI from your plan. There seem to be a surfeit of techniques, suggestions and tips for SEO. How do you separate the wheat from the chaff? Moreover, the importance of each option is closely linked to your own objective, and consequently your budget in terms of management time and expenditure. How do you measure ROI?

Americans conduct an estimated 790 million searches per week, according to research from ComScore Media Metrix. The Yankee Group reports that only 20 percent of U.S. companies tracking their search referrals are measuring performance beyond the initial click-through to the web site. This is very low. The key to a successful search engine marketing strategy is constant testing, revising and optimizing based on metrics. In order to maximize your return on investment, you must measure the performance beyond click-through—by measuring the complete interaction of your visitors with your site from acquisition to conversion to retention.

Search engine marketing is similar to advertising and other marketing campaigns. The marketing people would be interested in knowing the results of any campaign and most often the result is desired in terms of increase in company revenues. Name recognition, brand image creation and other outcomes are important; however the bottom line is sales.

ROI should be able to determine your website's conversion rate in terms of the ability to persuade your visitors to take the action that you desire them to take. At the same time, the success in terms of increase in sales depends on several other factors such as the competitiveness and value in your product or service offerings and the quality of your website to induce the visitor to take action. The search engine marketing effort is to bring the right targeted visitor, and bring the visitor to the right section at the right time. Converting such a targeted visitor to a customer is not within the ambit of SEO.

How does search engine optimization compare with other options for online advertising and promotion? Jupiter Research predicts that by 2006 the spending on E mail marketing shall be \$ 9.4 billion, whereas the spending on online advertisements shall be \$ 15.6 billion, and that on digital marketing initiatives such as campaigns, promotions, sweepstakes, coupons etc. shall be around \$ 19.3 billion. Forrester predicts that total online advertising spend should be \$ 42 billion by 2005 a significant 9.5% of total ad spending.

Costs in search engine marketing include the cost of optimization services as well as the paid inclusion and pay per click programs. Typically the cost of optimization for a small to mid-sized site through a firm may run from a few thousand dollars to over \$ 50000 for a large site. If you do it in-house consider around four person months' work plus around two person months of maintenance time over the ensuing six months.

How do you compare costs per click or costs per thousand impressions for search engine optimization vis-à-vis other forms of advertising? PwC report that the effective cost per thousand impressions or views (CPM) for an online ad is around \$ 3.50 for a general site (though it could be higher at \$ 10 to \$ 100 for other specialized sites) compared to \$ 19 for newspaper ads and \$ 16 for Prime-time TV. The CPM for search engines is certainly higher, but then it leads to targeted impressions. Thus it is important to look at targeting and conversion rates, not just CPM. Clickthrough rates for ad banners, in general, can be as low as 0.4 to 0.5%. Conversion rates would show you what percentage of visitors actually resulted in business. As discussed below, this may be easier done in e-commerce sites.

Cost per click for important keywords at Overture could be as high as \$ 0.50 to \$ 5 or \$ 10. The average bid reported according to one source is \$0.73 per click (for Overture). Thus this is already lower than the cost per visitor for online ad banners. Overture can only cater to a small percentage of search traffic, while Google, Yahoo and others comprise the larger portion; thus it can be said with certainty that SEO is the most cost effective ROI generator.

Balancing Paid Search Engines and programs with Free Search Engines for maximum ROI

It is crucial that you optimize your performance with both Paid programs as well as free listings. Search engine marketing is like playing 3-D chess—it is complex and multi-

dimensional. To win, the marketer must not only select the right keywords and align them with content that matches the user's query target, but also select the right search vehicle for each keyword.

The marketer must weigh the value of long-term results gained from organic search listings against the short-term returns of paid advertising. In most cases, the right answer is to have a balance of both.

It is a misconception that free listings do not generate traffic. In fact, the top 2 free listings receive more than 50% of the clicks. This is too high a number to ignore. You would be missing out on important traffic if you completely ignore free search engines. On the other hand, paid listing bring you focused traffic.

This again supports the fact that a balance of both kinds of listings is extremely vital for successful Search Engine Marketing.

ROI Tracking Tools

There are several tracking tools that measure the traffic coming to your website and that can identify what search engines and what keywords have brought that traffic. Moreover tracking tools can also tell you what the visitor has done online, including page visits, time spent, actions taken (relevant for e-commerce sites or query/contact forms) and so

on. You will have to find your survey methodologies to relate the traffic history to the revenue results. Some of these are mentioned below:

DoubleClick's DART

<http://www.doubleclick.com/us/advertisers/adserving/dart5/>

DoubleClick's DART paid placement and advertising ROI tracking tool generates web-based reports on cost per click, number of clicks, overall media cost, conversion rates, gross sales, return on investment (ROI), and net profit. Such tools are suited to online merchant sites.

WebTrends from NetIQ

<http://www.netiq.com>

Offers server-based software to determine how many visitors are coming to the Web site, where they're coming from, what they're doing on the Web site, and which search engines are sending the most traffic, along with which phrases drive the most traffic from each search engine.

HitBox from WebSideStory

<http://www.websidestory.com/>

It uses a different data collection method, by enabling site owners to add code to their site. It tracks the same type of information that the WebTrends product does.

Urchin

<http://www.urchin.com/>

This is another useful measuring and tracking tool.

11. Resources and Tools for Search Engine Marketing

The final section of this guide focuses on some of the most popular and effective tools as well as resources for the Search Engine Marketer. We discussed earlier, the importance of using various software for all aspects of Search Engine Marketing. This section lists out some of the very useful software tools.

It also highlights a few comprehensive sources for Search Engine News and Information.

11.1 Search Engine Marketing Professional Organization (SEMPO)

The Search Engine Marketing Professional Organization (SEMPO) was formed recently with the objective of increasing awareness and promoting the value of Search Engine Marketing Services. This is this biggest organization in its field comprising of more than 300 reputed Search Engine Marketing firms.

SEMPO's mission is to not only increase understanding of Search Engine Marketing amongst marketers but also to provide core educational materials about tools, vendors, consultants, programs and successful search engine marketing campaigns.

SEMPO represents the common interests of more than 300 search engine marketing professionals and provides them with a credible voice in the marketplace. To achieve its objectives, SEMPO will demonstrate how search engine marketing programs have become important components of the marketing mix.

SEMPO's website can be accessed at <http://www.sempo.org>

11.2 Tools

We studied a few ROI tracking tools in the earlier section. Listed below are some of the tools that cater to other aspects of Search Engine Marketing.

Alexa

Alexa, an Amazon.com company was founded in 1996 and provides various services ranging from Web Search and Site Stats to a load of other marketing tools for the webmaster.

The site can be accessed at <http://www.alexa.com>

Keyword success SEM tool suite

Keyword Success provides a host of tools for the Search Engine Marketer. These include tools for monitoring web pages with keyword alerts, keyword density analyzers, listing simulators to improve search engine rankings, and tools to measure success of keyword campaigns.

The site can be accessed at

<http://www.keywordsuccess.com/search-engine-optimization-tools.asp>

LinkPopularity.com

LinkPopularity.com runs a service that lets you know the link popularity of your website with most of the popular Search Engines. This is an extremely useful tool considering the importance of link popularity for Search Engine Marketing.

The site can be accessed at <http://www.linkpopularity.com>

Live Keyword Analysis

This tool as the name suggests analysis your website's keyword density. All you have to do is insert the insert your keywords and the text of your web page and automatically receive feedback about your keyword density.

This tool can be accessed at <http://www.live-keyword-analysis.com>

Tools from Marketleap

Marketleap offers a set of Search Engine Marketing and Analysis tools. These tools also give you critical data necessary to help you improve upon your current search engine marketing efforts.

Some of these are:

- Keyword Verification tools to check whether for a particular keyword your website would be ranked within the top 3 listings of Search Engines
- Search Engine Saturation tools that calculate the number of pages a given search engine has in its index for your website domain
- Link Popularity Check that gives a complete analysis of your link popularity.

The site can be accessed at <http://www.marketleap.com/services/freetools>

NetMechanic HTML Code checker

This tool checks the HTML code on your web pages for errors. Some of its functions include testing links, checking HTML code, rating page load time, and checking spelling.

The site can be accessed at <http://www.netmechanic.com/toolbox/html-code.htm>

OptiLink Link Reputation Analyzer

This tool analyzes your site to a top site for Search Engines. It provides recommendation on how to change the linking structure to improve the ranking of your website.

The tool can be accessed at <http://www.optitext.com/optilink/index.html>

PositionPro

Position Pro™ is a combination of tools providing you the ability to analyze your entire website like a search engine. It helps in finding pages that are acceptable to Search Engines, optimize keywords for content and phrases, and other tools for improving search engine visibility. This is a very useful tool.

The site can be accessed at <http://www.positionpro.com>

Website Management Tools

Website management tools.com offers a variety of solutions that analyze how well your website promotion strategy is working and where to focus your website promotion efforts for optimum results.

The tools include website ranking manager tools and robot manager tools and can be accessed at <http://www.websitemanagementtools.com/>

There are many other similar tools available on the Net. A few other tools worth checking out are:

- Robots.txt Validator - <http://www.searchengineworld.com/cgi-bin/robotcheck.cgi>
- Search Engine Optimizer - <http://www.se-optimizer.com/>
- SEM ROI Calculator - <http://www.pageviews.com/calculator.htm>
- SEOTool Set - http://www.bruceclay.com/web_tool.htm
- SiteSnar - <http://www.sitesnare.com/>
- WebPosition Gold - <http://www.webposition.com/>
- WebRank - <http://www.webrank.com/>
- WordTracker - <http://www.wordtracker.com/>

11.3 Search Engine Information Resources

- [About.com: Web Search](#)
- [Search Engine Watch](#)
- [Search Engine World](#)
- [Traffick](#)
- [up2speed](#)
- [Indicateur.com](#)

- [Internet Search Engine Database](#)
- [Media Post](#)
- [ClickZ](#)
- [DIRECT and Catalog Age's Search Engine Marketing site](#)
- [Google Dance Monitor](#)
- [Google Zeitgeist](#)
- [RankWrite Roundtable](#)
- [SearchEngineblog.com](#)
- [Search Engine Dictionary](#)
- [Search Engine Guide](#)
- [High Rankings Advisor](#)
- [W3 Search Engines List](#)
- [Web Trends](#)
- [WebMama Journal](#)
- [Yahoo! Buzz Index](#)
- [iMedia](#)
- [NetMechanic Newsletter](#)
- [Pandia Search Central](#)
- [Planet Ocean](#)
- [Search Engine Optimization Tips](#)
- [International Internet Marketing Association](#)
- [World Association of Internet Marketers](#)